MEKELLE UNIVERSITY

GRADUATE STUDIES PROGRAMME

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

SPATIAL DISTRIBUTION AND DETERMINANTS OF ANAEMIA AMONG WOMEN OF REPRODUCTIVE AGE IN ETHIOPIA USING MIXED-EFFECT ORDINAL LOGISTIC REGRESSION MODEL

BY:

SOLOMON TEKESTE TSEGAY

A THESIS SUBMITTED TO THE DEPARTMENT OF STATISTICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF SCIENCE IN BIOSTATISTICS

DECEMBER, 2024

MEKELLE, ETHIOPIA

MEKELLE UNIVERSITY

GRADUATE STUDIES PROGRAMME

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

SPATIAL DISTRIBUTION AND DETERMINANTS OF ANAEMIA AMONG WOMEN OF REPRODUCTIVE AGE IN ETHIOPIA USING MIXED-EFFECT ORDINAL LOGISTIC REGRESSION MODEL

BY:

SOLOMON TEKESTE TSEGAY

A THESIS SUBMITTED TO THE DEPARTMENT OF STATISTICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF SCIENCE IN BIOSTATISTICS

ADVISOR: SAID MUSSA (ASSISTANT PROFESSOR)

CO-ADVISOR: LEWTNESH BERIHUN (MSc.)

DECEMBER, 2024

MEKELLE, ETHIOPIA

ACCEPTANCE LETTER

This is to certify that the thesis conducted by **Solomon Tekeste Tsegay** entitled "Spatial Distribution and Determinants of Anaemia among Women of Reproductive Age in Ethiopia", is submitted in partial fulfillment of the requirement of the degree of Master of Science in Biostatistics complies with regulation of Mekelle University and meets the accepted standard concerning originality and quality.

Signed by Examining Committee:

External Examiner Yemane Hailu (Ph.D.) signature _____ date December 30, 2024

Internal Examiner Haftu Legesse (Ph.D Candidate) signature _____ date Dec 30, 2024

Advisor Said Mussa (Assistant Profesor) signature _____ date December 30, 2024

Lewtnesh Berihun (M.Sc in Biostatistics)

Chief of Department /Graduate Program Coordinator

Mekelle University

Acknowledgements

I would like to extend my profound gratitude to my advisor, Said Mussa (Assistant Professor), for his unconditional support, guidance, and encouragement during this study. I would like to express my heartfelt thanks to my co-advisor, Lewtnesh Berihun (MSc) for her encouragement, supervision, and constructive comments that develop my research abilities and in improving the quality of the Thesis. It has been a great honor and pleasure for me to work with both of you. I would like to extend my gratitude to Nebyu Mohammed (MSc) from Arsi University, Ethiopia, for being such a great person who always giving ideas and comments from different dimensions. You allowed me to be able to work under minimal supervision and at the same time not deviate from the objective of this thesis.

My immense gratitude goes to my family especially to my mother, Tsegay Gidey Aregay, for her prayers, encouragement, and for believing in me even during the moments when I thought it was impossible. Also, a heartfelt gratitude for my father for his kindness at all, **"This is my gift to my father, Tekeste Tsegay G/yesus for his memories"**, rest in peace my father.

Furthermore, I would like to express my deepest thanks and appreciation to the DHS team for allowing me to use their data to write my research study. Finally, I would like to say a big thanks to my friends and all the staff of the Statistics Department at Mekelle University for their support, kind words, and encouragement throughout the entire period.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CSA	Central Statistical Agency
DHS	Demographic Health Survey
DRC	Democratic Republic of Congo
EAs	Enumeration Areas
EMDHS	Ethiopia Mini Demographic and Heath Survey
EPHC	Ethiopia Population and Housing Census
FMoH	Federal Minstry of Health
GLMM	Generalized Linear Mixed Model
G/dL	Gram per Deciliter
Hb	Hemoglobin
HIV	Human Immune Deficiency Virus
ICF	International Commercial Fund
INLA	Integrated Nested Laplace Approximation
OLR	Ordinal Logistic Regression
ENP	Effective Number of Parameter
RBC	Red Blood Cell
SNNPR	South Nations Nationalities and Peoples Republic
SSA	Sub-Saharan Africa
WHO	World Health Organization
WRA	Women of Reproductive Age

Abstract

Background: *Anaemia is a condition characterized by a low blood hemoglobin concentration (120 g/L in non-pregnant women and below 110 g/L in pregnant women). It primarily affects women of reproductive age (WRA) and who suffer from anaemia has experienced detrimental effects on their mental development and future social functioning.*

Objectives: *This study aimed to assess the spatial distribution and determinants of anaemia among women of reproductive age in Ethiopia.*

Methods: *The study participants were all the WRA who were confirmed to anaemia from the 2019 EMDHS data source. The survey considered 8885 WRA; of which 1483 severe, 534 moderate, 1778 mild and 5090 none anaemia levels were included in this study. The study variable was defined as the ordinal level of anaemia (none, mild, moderate and severe) based on the WHO cut-off points. In this study, Moran's-I, was used to investigate the presence of spatial autocorrelation. A mixed effect ordinal logistic regression model used allowed to analyze random and fixed effects of some covariates, spatial effects, and other fixed covariates. Inference used a full GLMM and several methods can be used to assess the goodness of fit in GLMMs, including the AIC and BIC techniques.*

Results*: Out of 8885 WRA; included in this study 1483(16.7%) were found at a severe level of anemic. Due to the BIC model selection criteria, the GLMM model was found to be appropriate. From the model Individual factors (age group, religion, wealth index and marital status) and community factors (cooking fuel type, number of children, access to electricity, and having refrigerators) are found to be determinants, significant determinants of anaemia status among WRA and the spatial analysis demonstrated a clustered pattern of anaemia distribution, confirmed by the global Moran's I statistic (0.146652, p-value < 0.001).*

Conclusion: *The finding revealed a spatial variation of anaemia status among WRA across the regions of Ethiopia with higher prevalence in the eastern parts of the country specifically in Somali and Harari regions. The application of the GLMM provided a more detailed understanding of the factors influencing anaemia status.*

Keywords: Anaemia status; GLMM; Brant test; BIC; Spatial distribution, Spatial autocorrelation.

CHAPTER ONE

1. INTRODUCTION

1.1. Background of the study

Anaemia is a condition characterized by a low blood hemoglobin concentration (120 g/L in non-pregnant women and below 110 g/L in pregnant women). In such conditions, the oxygen-carrying capacity of Red blood cells (RBCs) is insufficient to meet physiological needs. It is a global public health problem, particularly in low- and middle-income countries (Czempik and Wiórek, 2020). While more recent global data for 2023 might not be readily available now due to the lag in global health data reporting and analysis, anaemia remains a significant public health issue affecting both children and women worldwide, particularly in low and middle-income countries. Anaemia increases the risk of infection and death impairs cognitive performance and affects 571 million women of reproductive age (Sappani et al., 2023).

About one-fourth of the world's population suffers from anaemia, a global public health issue that affects both developed and developing nations and has significant effects on human health and socioeconomic development. It primarily affects women of reproductive age and children under five (Gemechu et al. 2023, Gebreegziabher and Sidibe, 2023). Women of reproductive age who suffer from anaemia may experience detrimental effects on their mental development and future social functioning. Iron deficiency anaemia in children during the first two years of life is associated with delayed cognitive development, poorer academic performance, and a reduced ability to work later in life (Yadav and Nilima, 2021). In addition to being a leading cause of prenatal and postpartum morbidities and mortality among women in developing nations, it also has an impact on the physical and cognitive development of children. For this reason, the health of women in their reproductive years, in general, is crucial to both their survival and the nation's future progress (Ayele et al., 2019). Infectious diseases including malaria and hookworm infections, genetic abnormalities, and nutritional deficits can all lead to anaemia. The most frequent cause, which accounts for 50% of cases in high-income countries, is iron insufficiency (Ayele et al., 2019). In highly malaria-endemic countries, particularly in Sub-Saharan Africa (SSA), malaria is a significant contributing factor to women's reproductive age anaemia (Lingani et al., 2024).

Anaemia affects a substantial portion of the population globally, with major impacts on public health and economic productivity. The burden is highest among preschool-aged children and women of reproductive age, particularly pregnant women. Anaemia is a common health problem among women of reproductive age (15-49) worldwide. It is estimated that around 32% of women in this age group suffer from anaemia globally (Alem et al., 2023). Globally, anaemia is recognized as a condition that significantly hampers economic progress and impedes the empowerment of women. It is most prevalent in developing countries where poor nutrition and inadequate healthcare systems, combined with infectious diseases such as malaria and HIV/AIDS, exacerbate the incidence. According to the WHO, anaemia contributes to 20% of all maternal deaths. The global response to anaemia involves a multifaceted approach, including iron supplementation, improved dietary diversity, and fortified foods, as well as interventions targeting infectious diseases. Anaemia is a condition in which the body lacks enough healthy red blood cells to carry adequate oxygen to the body's tissues. This can lead to fatigue, weakness and other health complications. There are several causes of anaemia among reproductive age of women, including iron deficiency, vitamin B12 deficiency, chronic diseases, blood loss, poor diet anaemia, etc. The basic and important function of iron (as a part of hemoglobin) is to transport oxygen from lungs to cells in our body and is an essential requirement of the body (Mawani and Aziz Ali, 2016). It is a severe public health problem in many countries of the world. As a global public health problem, anaemia impairs women's health and well-being and increases the risk of adverse maternal and newborn outcomes in low and middle-income countries affecting half a billion women of reproductive age worldwide. Defined as a condition characterized by a lower-than-normal concentration of hemoglobin(Hb) (Abioye et al. 2024), anaemia contributes to maternal and parental mortality. It doubles the risk of maternal death and accounts for 20% of all maternal deaths (Safiri et al., 2021). Worldwide, each year, anaemia results in more than 115,000 maternal and 591,000 parental deaths (Zelelew, 2012).

Africa has one of the highest burdens of anaemia worldwide, driven by factors such as malnutrition, infectious diseases including malaria and helminthiasis (parasitic worms), and a lack of access to healthcare services. The distribution can vary widely between countries and within regions, often exacerbated by socioeconomic disparities. Africa was ranked the second highest in terms of anaemia (38.6%) and severe anaemia (1.8%; World Health Organization, 2015). It is a moderate and severe public health problem in all WRA (38.6%) and pregnant

women (54%) of the Sub-Saharan Africa (SSA), respectively (Alem et al., 2023). Within Africa, anaemia is particularly widespread, influenced by a confluence of factors including poverty, food insecurity, and infectious diseases like malaria, which is endemic in many parts of the continent (Semedo et al., 2014). Malnutrition, resulting from inadequate intake of iron, folic acid, vitamin B12, and other essential nutrients, is a major driver of anaemia among WRA. The African Union, through its Agenda 2063 and the adoption of the Continental Nutrition Strategy for Africa 2015-2025, underscores the need for concerted efforts to improve nutrition and decrease the prevalence of anemia(Aglago et al., 2018). Anaemia distribution in Ethiopia is greater (said to be 24%), with the highest rates found in the regional states of Afar and Somalia, where they were assessed to be around 44.7% and 59.5%, respectively (Tsegaye et al, 2021).

In Ethiopia, anaemia is a major public health concern, especially among women and children. Contributing factors include dietary deficiencies (such as iron deficiency), parasitic infections, genetic conditions like thalassemia (inherited disorders in which the amount of hemoglobin in the blood is reduced) sickle cell disease, and other micronutrient deficiencies(Andersen et al., 2022). In Ethiopia, the prevalence of anaemia for women of reproductive age ranges from 34% to 68.5% (Negash and Ayalew, 2023). The national data from the Ethiopian Demographic and Health Survey (EDHS) in 2011 given that the prevalence of anaemia among women of reproductive age was 44%, it was around more than four out of ten women of reproductive age were anaemic (Geta et al., 2022). The Ethiopian Federal Ministry of Health (FMoH) has been struggling to prevent anaemia by focusing on the reproductive age of women by providing iron (Fe) and folic acid, nutrition education, providing drugs for deworming, promoting sanitation and preventing and treating malaria (Tsegaye et al., 2021). In Ethiopia, anaemia remains a significant public health problem among WRA. The Ethiopia Mini Demographic and Health Survey (EMDHS) of 2019 serves as a critical data source that highlights the severity and pervasiveness of anaemia within this demographic (Sahile et a, 2022). The survey encompasses a wide-ranging dataset that provides insights into health, nutrition, and demographic indicators across various regions. According to the EMDHS 2019, anaemia in Ethiopian WRA has multi-factorial causes, including but not limited to nutritional deficiencies, infectious diseases, and cultural practices (Andersen et al., 2022).

However, in the last 15 years, the trend of anaemia has remained inconsistent or dynamic. Although the distribution of all forms of anaemia decreased from 27% in 2005 (Tsegaye et al., 2021) to 17% in 2011, it increased to 23.63% in 2016. The distribution of anaemia was higher among pregnant, breastfeeding, rural women and women who have had six or more births. It was also more prevalent with decreasing educational levels and household wealth index. Recognizing it as a worldwide public health problem, the WHO target is set to reduce anaemia in WRA by 50% in 2025 (Hasan et al., 2022).

It is important to note that anaemia is a significant public health concern, particularly among women of reproductive age, as it can have adverse effects on maternal and child health outcomes. Various factors contribute to the prevalence of anaemia, including nutritional deficiencies, chronic diseases, and socioeconomic factors (Alem et al., 2023). Public health interventions and strategies aimed at improving nutrition, access to health care and addressing underlying causes are crucial in reducing the prevalence of anaemia among reproductive age women (World Health Organization Team, 2014).

The study aimed to examine the spatial distribution and identify determinants of anaemia in Ethiopia using the EMDHS2019 data. First, the study examines whether there is a significant global spatial autocorrelation for the prevalence of anaemia. When the existence of global spatial dependency is confirmed, the subsequent objective be to explore local spatial autocorrelation, and map the spatial distribution of anaemia prevalence by demographic, geographic, and socioeconomic factors. The generated anaemia prevalence map would have important implications for targeting policy for better intervention, and identifying variables that might account for the observed spatial distributions.

1.2. Statement of the Problem

Despite extensive global efforts to enhance maternal health, anaemia remains a significant public health concern, particularly affecting women of reproductive age. The 2019 Ethiopia Mini Demographic and Health Survey (EMDHS 2019) highlighted the enduring prevalence of anaemia among this demographic, potentially jeopardizing maternal and child health outcomes, diminishing quality of life, and impeding socioeconomic progress. However, the distribution and severity of anaemia among women of reproductive age exhibit variations across populations and

geographical regions, underscoring the need for a deeper understanding of the spatial patterns and diverse risk factors driving this health issue.

Conventional analytical methods often overlook the hierarchical and spatially correlated nature of health data, leading to biased estimations and ineffective intervention strategies. Utilizing mixed-effect ordinal logistic regression presents a more sophisticated methodological approach, facilitating the exploration of demographic, geographic, and socioeconomic determinants of anaemia while accommodating the ordered nature of anaemia severity and unobserved complexities within the dataset.

Previous researches on anaemia among women of reproductive age, utilizing spatial distribution and mixed-effect ordinal logistic regression (Woldu et al., 2020), have inadequately delved into the spatial variability of anaemia. There exists a potential gap in leveraging mixed-effect ordinal logistic regression models to comprehensively assess the demographic, geographic, and socioeconomic factors influencing anaemia. This study aims to bridge these gaps by conducting spatial analyses and applying mixed-effect ordinal logistic models to better understand and account for the multifaceted determinants of anaemia among women of reproductive age. The research questions endeavor to address three pivotal questions:

- ✓ What is the spatial distribution of anaemia among women of reproductive age in Ethiopia?
- ✓ Which factors exhibit significant associations with anaemia among women of reproductive age, drawing on the data from the EMDHS 2019?
- ✓ How can an anaemia prevalence map be generated for women of reproductive age in Ethiopia to better inform policymakers for targeted interventions and to identify variables responsible for the observed spatial disparities?

1.3. Objectives of the Study

1.3.1. General Objective

The general objective of this study was to assess the spatial distribution and determinants of anaemia among women of reproductive age in Ethiopia using a mixed-effect ordinal logistic regression model by using EMDHS 2019 data.

1.3.2. Specific Objectives

- To determine the spatial distribution of anaemia among women of reproductive age in Ethiopia.
- To examine the risk of factors that exhibit significant associations with anaemia among women of reproductive age, drawing on the data from the EMDHS 2019.
- To map the risk of anaemia among women of reproductive age across different regions in Ethiopia to identify potential hotspots and cold spots.

1.4. Significance of the Study

The study on spatial pattern and mixed-effect ordinal logistic regression of anaemia among reproductive-age of women using EMDHS 2019 data is significant for several reasons and the study can help identify areas with a high risk of anaemia among reproductive-age women. The study's findings may inform policies and programs aimed at reducing the burden of anaemia, which can improve the health outcomes of women. The significance of this study is multifaceted, reflecting its potential impact on health policy research and society particularly within the context of Ethiopia. Understanding the spatial pattern of anaemia can guide health policymakers in allocating resources more efficiently. Regions identified as hotspots for anaemia can be targeted for intensified interventions, such as nutritional programs or health education campaigns. By identifying the demographic, geographic, and socio-economic factors associated with anaemia, the study can inform the design of public health interventions that are tailored to the specific needs of different groups. This targeted approach can potentially lead to more effective management and reduction of anaemia.

1.5. Scope of the study

The scope of a study on the spatial distribution and mixed effect ordinal logistic regression of anaemia among women of reproductive age using EMDHS 2019 data can be defined along

several dimensions. The study has covered the extent of Ethiopia, as represented in the EMDHS 2019 dataset. It will consider regional variations within the country and focus on the spatial distribution of anaemia across these different geographic areas. The target population for the study is women of reproductive age within Ethiopia. The analysis has been limited to this group due to its particular vulnerability to anaemia and its implications for maternal and child health. The study is cross-sectional, based on the EMDHS 2019 data. It has provided a snapshot of anaemia prevalence and its correlates as captured at the time of the survey, without considering temporal changes or trends. The study has included spatial analysis to identify distributions and clusters of anaemia risk geographically. It has also employed mixed-effect ordinal logistic regression to assess the relationship between various demographic, geographic, and socio-economic factors and the severity of anaemia, accounting for the ordered nature of the anaemia. The analysis has been conducted using the variables and data points available in the EMDHS 2019 dataset. It relies on the recorded measures of hemoglobin levels, demographic factors, health indicators, and any geospatial information provided. The methodology of this study has focused on spatial analysis and mixed-effect ordinal logistic regression for ordinal outcomes. By establishing a clear scope for the study, the study has maintained focus and ensures that the study's design aligns with its objectives and available resources. The scope also sets the boundaries for the potential impact and generalizability of the study's findings.

1.6. Organization of the study

This thesis has been organized into sections that include an introduction; a literature review section that provides a review done by other authors about anaemia and assessing the relationship between influential and determinant factors of anaemia among reproductive-age of women. The methodology section has a detailed description of the study area, data source and sample design, sampling frame, sampling technique of the study, variables under the study explanatory variables of the study, spatial analysis and mixed-effect ordinal logistic regression. Finally, the analysis, result and discussion of this thesis have explained with simple and clear interpretation.

CHAPTER TWO

2. LITERATURE REVIEW

This section provides a review done by other authors about anaemia. The review discusses different statistical approaches previously used in assessing the relationship between influential and determinant factors of anaemia among the reproductive age of women.

2.1. Definition and Causes of Anaemia

Anaemia is a condition in which the number of red blood cells (and consequently their oxygen-carrying capacity) is insufficient to meet the body's physiologic needs, which in turn affects how well the cells can deliver oxygen (Chaparro, 2019). Age, gender, residence elevation above sea level (altitude), smoking habits, and various phases of pregnancy all affect a person's specific physiological requirements. The most common cause of anaemia worldwide is believed to be iron deficiency. However, anemia can also be caused by other nutritional deficiencies (such as those related to folate, vitamin B12, and vitamin A), acute and chronic inflammation, parasitic infections, and inherited or acquired disorders that affect the synthesis of hemoglobin, the production of red blood cells, or the survival of red blood cells. It is impossible to identify iron insufficiency just based on hemoglobin concentration. Even if iron deficiency is not the only cause of anaemia, it is nevertheless important to assess the hemoglobin concentration. The hemoglobin concentration can reveal information about the degree of iron shortage when combined with other tests of iron status, making the prevalence of anaemia a significant health indicator (Hoenemann et al., 2021)

2.1.1. Hemoglobin Threshold

Age, sex, physiological state, pregnancy status, and other factors affect normal Hb distributions. Pregnant women have lower hemoglobin levels than non-pregnant women (Chaparro and Suchdev, 2019). WHO Hb thresholds were used to classify individuals living at sea level as anaemic (Table 2.1) (WHO, 2022). Since smoking and altitude have been shown to affect hemoglobin distributions statistically and physiologically, the prevalence of anaemia corrected for these variables was used when it was provided by the survey (Sharma et al., 2019).

Table 2.1: Hb thresholds used to define anaemia

Groups	Hemoglobin threshold(g/l)
Children (0.50–4.99 yrs)	110
Children (5.00–11.99 yrs)	115
Children (12.00–14.99 yrs)	120
Non-pregnant women (≥15.00 yrs)	120
Pregnant women	110
Men (≥15.00 yrs)	130

Source: (WHO, 2022)

2.2. Burden of Anaemia among WRA in Africa and Globe

2.2.1. Global Burden of Anaemia among Women of Reproductive Age

Anaemia is a significant public health concern globally, particularly among women of reproductive age (Owais et al., 2021). According to the World Health Organization (WHO), anaemia affects around 32.8% of women of reproductive age worldwide (McLean et al., 2009). The main causes of anaemia in this population group are often related to nutritional deficiencies, such as iron, folate, and vitamin B12 (Williams et al., 2023).

2.2.2. Burden of Anaemia in Sub-Saharan African Countries

In sub-Saharan Africa, anaemia is a prevalent health issue among women of reproductive age due to various factors like poor diet, infectious diseases, and limited access to healthcare (Teshale et al., 2020). Data from a study by the WHO indicated that the prevalence of anaemia among women of reproductive age in sub-Saharan Africa is around 45% (Alem et al., 2023). The highest prevalence rates are often found in countries with high rates of poverty, limited healthcare infrastructure, and ongoing conflicts (Douthit and Alemu 2016). Anaemia prevalence rates can vary within sub-Saharan Africa based on factors like geographical location, socioeconomic status, and access to healthcare (Tesema et al. 2021). Some studies have shown that in certain countries in East Africa, anaemia prevalence among women of reproductive age can exceed 50%, while in Southern Africa, the rates may be slightly lower but still significant at around 40% (Teshale et al., 2020 ,Geta et al., 2022).

2.3. Empirical Literature Review

According to the result obtained from WHO, the prevalence of anaemia is high in developing countries due to the socioeconomic and health development. Africa and South East Asia countries are highly affected by anaemia (WHO, 2021).

Sub-Saharan African countries have a different iron deficiency anaemia epidemiology than other regions of the world (Mwangi et al., 2021). People at high risk of iron deficiency anaemia are more common in this region than in other regions of the world due to a combination of factors such as high infection rates, low-quality foods that are a poor source of iron, and demographic traits (Sharma et al., 2024). Here we evaluate the existing strategies and measures for the control of iron deficiency anaemia, as well as its causes, disease burden, and implications in the general population of this region. Concerns concerning the safety of iron interventions specifically, how it affects malaria and other infectious diseases dominate the present discourse (Raiten, Namasté, and Brabin, 2011). On the other hand, ubiquitous prenatal iron supplementation and delayed cord clamping are safe therapies that have the potential to enhance the health of both mothers and infants (Georgieff, Krebs, and Cusick, 2019). Safe and successful iron treatments for children require effective infection control. Along with strategies to improve iron consumption, more focus should be placed on methods intended to decrease iron loss from helminthes infections and menstruation (Mwangi et al., 2021).

Using EDHS 2005 data, the study by Umeta et al. (2008) has shown that women's educational status, grouped altitude of residential places and household wealth index categories have a significant impact on the prevalence of anaemia. The frequency of anaemia was positively correlated with the reproductive rate over the previous five years. The prevalence of anaemia in women was also linked to living in a remote region, not utilizing contraceptive methods, and lacking access to restroom facilities (Gebrerufael and Hagos, 2023).

The study conducted by (Ejigu, Wencheko, and Berhane, 2018) spatial patterns and determinants of anaemia in Ethiopia: in this study the analysis included complete cases from 15,909 females, and 13,903 males. Among all subjects who agreed to the hemoglobin test, 5.22% of males, and 16.60% of females were anemic. In both binary and ordinal outcome modeling approaches, educational level, age, wealth index, BMI, and HIV status were found to be significant predictors of anaemia prevalence. Determinants of anaemia in both multilevel mixed-effects binary and

ordinal logistic regression model at a 5% level of significance and as revealed by Moran's I test, significant spatial autocorrelation was noted across clusters.

The region of the women is significantly associated with an increased risk of infections, anaemia status among reproductive-age women. In the study conducted by (Kibret et al., 2019) spatial distribution and determinant factors of anaemia among women of reproductive age in Ethiopia: a multilevel and spatial analysis, an in-depth analysis of the 2016 EDHS data was undertaken. In this study, it was found that there is a variation of the anaemia rate in terms of education status of the women. A higher proportion of anaemic cases were observed among women with no education. It was found that women who did not have formal education had higher odds of anaemia than those with higher education. In addition, in this study a higher proportion of anaemic cases were observed among women in the poorest wealth quintiles. The lowest wealth quantile compared with the highest quantile was associated with a higher risk of anaemia. Results of this study show that women who were in the poorest wealth quintile were 30% more likely to be anaemic than women who belong to the richest quintile.

The study conducted by Josephine (Nti et al., 2021)the data from the DHS (2014, 2016, 2014-2015, 2015-2016, and 2016) from Ghana, Ethiopia, Rwanda, Tanzania, and Uganda, respectively, their study result showed that factors that significantly increased the risk of anaemia among the selected countries. Their result shows that factors like marital status significantly increased the rate ratio of a woman-getting anaemia across all five countries. Different factors increased the rate ratio of a woman-developing anemia in some countries. In Ghana, the wealth index was factor that increased a woman's likelihood of getting anaemia. In Rwanda, marital status and contraceptive use statistically significantly increased the woman's rate ratio of developing anaemia. When all five countries were put together, the partner's occupation, marital status, and contraceptive use increased the rate ratio of woman-developing anaemia.

According to the available literature, the place of residence has been identified to be significantly associated with anaemia status of women of reproductive-age. The study conducted by (Hounton et al., 2015) applying multilevel ordinal logistic regression models, 2016 Ethiopia Demographic and Health Survey (EDHS) data set collected from nine regions and two administrative cities of Ethiopia; their result shows that the wealth index of the household was positively associated with anaemia in women, rural women had greater odds of higher levels of anaemia as compared with

urban women, women who were living in Somali, Afar and Harari regions and Dire Dawa city had greater odds of higher levels of anaemia as compared with women living in Tigray region.

The marital status of women of reproductive-age is significantly associated with anaemia status. In the study conducted by (Tsegaye et al., 2021) bivariate and multivariable logistic regression analyses were done to identify predictors of the trend of anaemia among women; the data from three consecutive Ethiopian Demographic and Health Survey (EDHS) from 2005–2016 were analyzed in this study; in this study, their result shows that single women had a lower chance of getting anaemia than their counterparts and this study revealed that widowed and separated women had higher odds of having anaemia than their counterparts. This might be explained by the fact that widows and women who have been separated from their husbands are vulnerable to economic deprivation, hunger, starvation, and a lack of access to health care.

The impact of the nutritional situation of women, measured using the Body Mass Index (BMI**,** defined as the weight in kg divided by the square of height in meters) on the women's anaemia status, BMI is significantly associated with anaemia status. In the study conducted by (Gautam et al., 2019) using complex sample logistic regression, the association between dependent and independent variables was examined using the data 2016 Nepal Demographic and Health Survey (NDHS). This study also found that more than fifty percent of women were <20 years old when they gave their first birth and the rate of anaemia among them was also higher. Therefore, another possible reason for the high prevalence of anaemia among younger women could be due to teenage pregnancy. This study showed a high prevalence of anaemia among underweight women. Similarly, overweight women were less likely to be anaemic as compared with women with normal BMI. The prevalence of anaemia was found to be lower among married women who independently decided her healthcare. This can be explained by the women's relative position in the household in terms of decision-making.

The educational level of women is significantly associated with almost all studies of the anaemia status of reproductive-age of women. In the study conducted by (Woldu et al., 2020) both bivariable and multivariable binary logistic regression models were fitted to identify associated factors of anaemia, a community-based cross-sectional study was conducted from February to April among 359 reproductive-aged women (WRA). In this study, educational status was found to be significantly associated with the prevalence of anaemia among study participant's women

of reproductive age with no formal education were more likely to be anaemic as compared to WRA having secondary and above educational status. The odds of to be anaemic in this study were 2.28 times higher in women with no formal education than in women attending secondary were and above educational status. The age category of the study participants was also found to be significantly associated with the prevalence of anaemia is study revealed that being in the age category of 36 and above is 2.64 times more likely to be anaemic as compared to the age categories of 15–24 years. In this study, the prevalence of anaemia among reproductive age women was significantly associated with body mass index.

The current age of the women has been identified to be significantly associated with anaemic status of the reproductive age of women. In the study conducted by (Wu et al., 2020) the association between the prevalence of anaemia and sociodemographic characteristics of pregnant and non-pregnant women was analyzed using univariate and multivariate logistic regression methods and a longitudinal observational study, which involved 640,672 women aged 18–49 years from 129 counties in southwest China. In this study, the prevalence of anaemia was significantly higher in women aged 18–20 years and over 35 years than in women aged 25–30 (under 20 years: adjusted odds ratio = 1.28; over 35 years: adjusted odds ratio = 1.07). When women of reproductive age are too young, risk factors include incomplete development of organs and tissues, unmarried status, low level of education, poor financial conditions, and low body weight, which can lead to notable increases in the incidence of anaemia.

Some of the socioeconomic factors that have been studied as indicators of the socio-economic status of a household were the household income/wealth index. In the study conducted by (Kibret et al., 2019) spatial analysis and a multilevel logistic regression model were used to identify independent predictors of anaemia among women of reproductive age. In this study, women aged 40–49 years had a lower likelihood of being anaemic compared with women aged between 15 years and 19 years. It was found that there is a variation in the anaemia rate in terms of the education status of the women. A higher proportion of anaemic cases were observed among women with no education. It was found that women who did not have formal education had higher odds of anaemia than those with higher education. A higher proportion of anaemic cases were observed among women in the poorest wealth quantile. The lowest wealth quantile compared with the highest quantile was associated with a higher risk of anaemia. This study

revealed there to be a significant difference in the proportion of anaemic cases according to the place of residence (urban/rural). The likelihood of having anaemia was higher for rural residents compared with urban residents.

The study conducted by (Agegenehu et al., 2021) spatial distribution and determinants of iron supplementation among pregnant women in Ethiopia: a spatial and multilevel analysis, spatial and multilevel analysis used. In this study, the spatial analysis revealed that the spatial distribution of iron supplementation among reproductive-age women was significantly varied across the country. In this study, women's education is one of the important factors for iron supplementation among pregnant women. The chance of getting an iron tablet is higher among women with primary and higher education as compared to no education. The region is also significantly associated with iron supplementation among pregnant women. Those pregnant women from Tigray, Afar, Amhara, and Dire Dawa regions had a higher chance of taking the iron tablet as compared to the Somali region. The household wealth index is another important factor in iron supplementation among pregnant women. Mothers who were in poorer, middle, and richer wealth quantile categories higher chance of taking iron supplementation as compared to the poorest wealth quantile.

According to the available literature, toilet facilities of households have been identified to be significantly associated with the anaemia status of women of reproductive age. The study conducted by (Gebrerufael and Hagos, 2023) Anaemia Prevalence and Risk Factors in Two of Ethiopia's Most Anaemic Regions among Women: an ordinal logistic regression model was taken into consideration and a cross-sectional study, in this study anaemia is a worldwide problem that affects all individual groups. The two populations that are most susceptible to anaemia are women and children. Women who live in the Somali region are more likely to have anaemia than women who live in the Afar region. The inaccessibility and unavailability of healthcare facilities may be the cause of the increased prevalence of anaemia among women from the Somali region. As a result, they are ignorant of the risk factors contributing to anaemia and the available preventative measures. Similarly, to this, pregnant women had higher anaemia levels than non-pregnant women did. The results of this study also showed that women were more likely to be anaemic in families with unimproved toilet facilities than in households with improved toilet facilities.

The study conducted by (Tusa, Kebede, and Weldesenbet, 2021) Spatial distribution and determinant factors of anaemia among adults aged 15– 59 in Ethiopia; using mixed-effects ordinal logistic regression model the spatial analysis result showed that the spatial distribution of anaemia among adults was significantly varied across the country. In multivariable mixed-effect ordinal regression analysis; sex, marital status, educational level, place of residence, region, wealth index, and body mass index were significant predictors of the level of anaemia among adults in Ethiopia. The present study documented that the spatial distribution of anaemia among adults significantly varied across the country; significant hotspot areas of anaemia were identified in the Somalia, Afar, Gambela, Dire Dawa, and Harari regions. The present study documented that the odds of rural adults experiencing severe anaemia (against moderate, mild, or non-anemic) were higher than the odds of urban adults. The possible reasons might be rural adults are more likely to have low socio-economic status, a low chance of accessing iron-rich foods, and lack of adequate nutrition information as compared to urban adults.

CHAPTER THREE

3. METHODOLOGY

3.1. Description of the Study Area

This study was conducted in Ethiopia. Ethiopia is strategically located in the northeastern part of Africa popularly known as “The Horn of Africa". It shares a boundary with North and South Sudan on the west, Somalia and Djibouti on the East, Eritrea on the North and northwest, and Kenya on the South. Ethiopia officially known as the Federal Democratic Republic of Ethiopia, is a landlocked country located in the Horn of Africa (Takele and Tolcha, 2021). It is the second-most populous nation in Africa, with over 126.5 million populations and the tenth largest by area, occupying 1,100,000km^2. Ethiopia has fourteen geographic or administrative regions: twelve regional states (Tigray, Afar, Amhara, Oromia, Somali, Benishangul-Gumuz, Gambela, Harari, Sidama, South-west Ethiopia, South Ethiopia and Central Ethiopia) and two city administrations (Addis Ababa and Dire Dawa that are considered as a region) with a capital city of Addis Ababa.

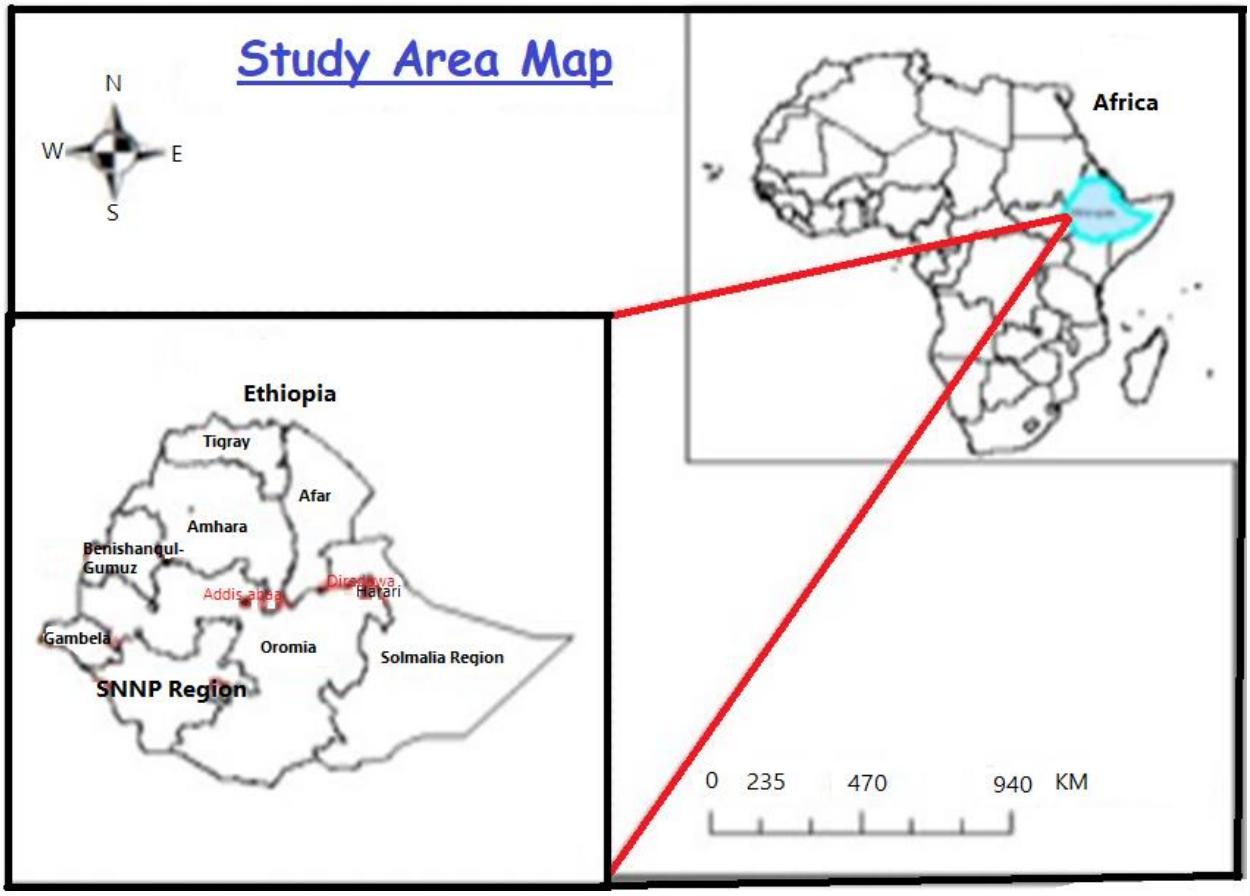


Figure 3.1: Map of Study Area

3.2. Data Source and Sample Design

The 2019 Ethiopia Mini Demographic and Health Survey (EMDHS) is the second Mini Demographic and Health Survey data conducted in Ethiopia. The Ethiopian Public Health Institute (EPHI) implemented the survey at the request of the Federal Ministry of Health (FMoH). Data collection took place from March 21, 2019, to June 28, 2019. In this study, the EMDHS 2019 data were downloaded from the DHS website https://dhsprogram.com after being granted permission. More specifically the EMDHS conducted hemoglobin testing on the reproductive age of women on the prevalence of anaemia in these groups. Women of reproductive age who stayed in the household and were eligible to be measured for hemoglobin during the survey comprise the study population, whereas the target population of this study was the reproductive age of women in Ethiopia.

3.3. Sampling Frame

The sampling frame used for the 2019 EMDHS is a frame of all census enumeration areas (EAs) created for the 2019 Ethiopia Population and Housing Census (EPHC) and conducted by the Central Statistical Agency (CSA). The census frame is a complete list of the 149,093 EAs created for the 2019 EPHC. An EA is a geographic area covering an average of 131 households. The sampling frame contains information about EA location, type of residence (urban or rural), and estimated number of residential households. The sample for the 2019 EMDHS was designed to provide estimates of key indicators for the country as a whole, for urban and rural areas separately, and each of the nine regions and the two administrative cities.

3.4. Sampling Technique of the Study

The 2019 EMDHS sample was stratified and selected in two stages. Each region was stratified into urban and rural areas, yielding 21 sampling strata. To ensure that survey precision was comparable across regions, sample allocation was done through an equal allocation where 25 EAs were selected from eight regions. However, 35 EAs were selected from each of the three larger regions: Amhara, Oromia, and the Southern Nations, Nationalities, and Peoples' Region (SNNPR). In the first stage, a total of 305 EAs (93 in urban areas and 212 in rural areas) were selected with probability proportional to EA size (based on the 2019 EPHC frame) and with independent selection in each sampling stratum. Some of the selected EAs for the 2019 EMDHS were large, with more than 300 households. In the second stage of selection, a fixed number of 30 households per cluster were selected with an equal probability of systematic selection from the newly created household listing. All women aged 15-49 who were either permanent residents of the selected households or visitors who slept in the household the night before the survey were eligible to be interviewed. In all selected households, height and weight measurements were collected from children aged 0-59 months, and women aged 15-49 were interviewed using the Woman's Questionnaire.

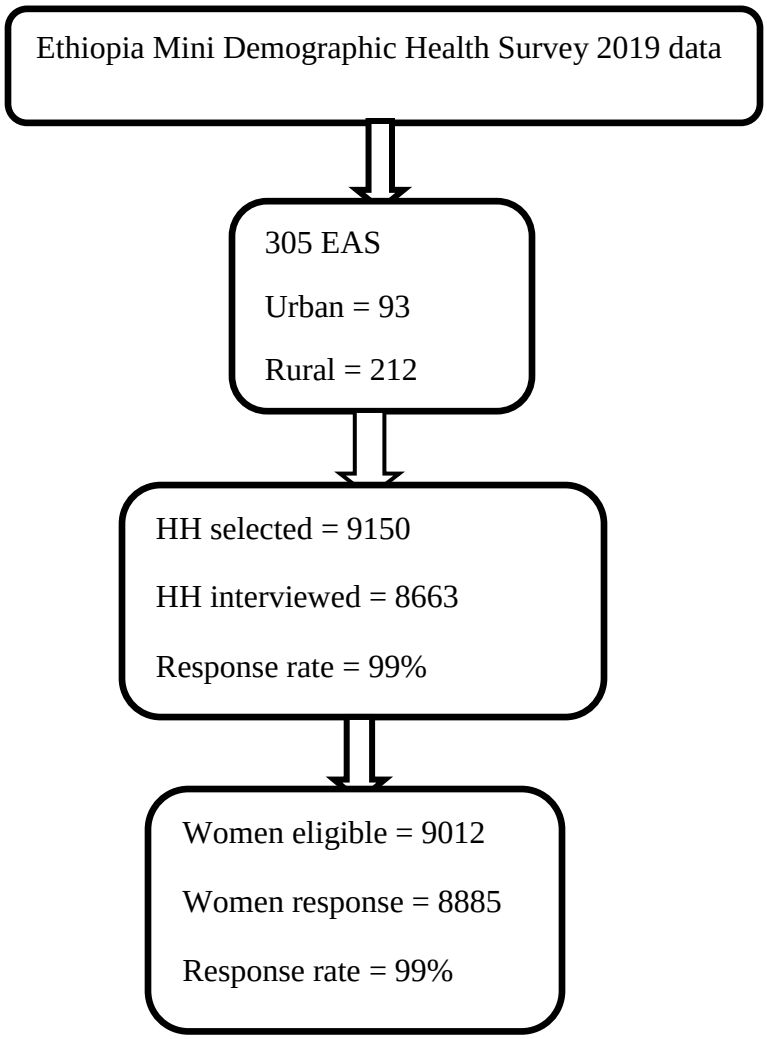


Figure 3.2: Framework for sampling technique of EMDHS2019 data among reproductive age of women

3.5. Variables under the Study

Anaemia status in this study is defined by the concentration of hemoglobin in the blood, a measure that indicates the capacity of the blood to carry oxygen to the body's tissues. The EMDHS 2019 data includes hemoglobin levels measured from blood samples taken from the study participants. The measurement was typically performed using a standardized method such as a finger prick test conducted with a portable hemoglobin meter, which provides immediate results.

The World Health Organization (WHO) provides widely accepted thresholds for defining anaemia, which is adjusted for age, gender, altitude, smoking status, and pregnancy status (Putra,

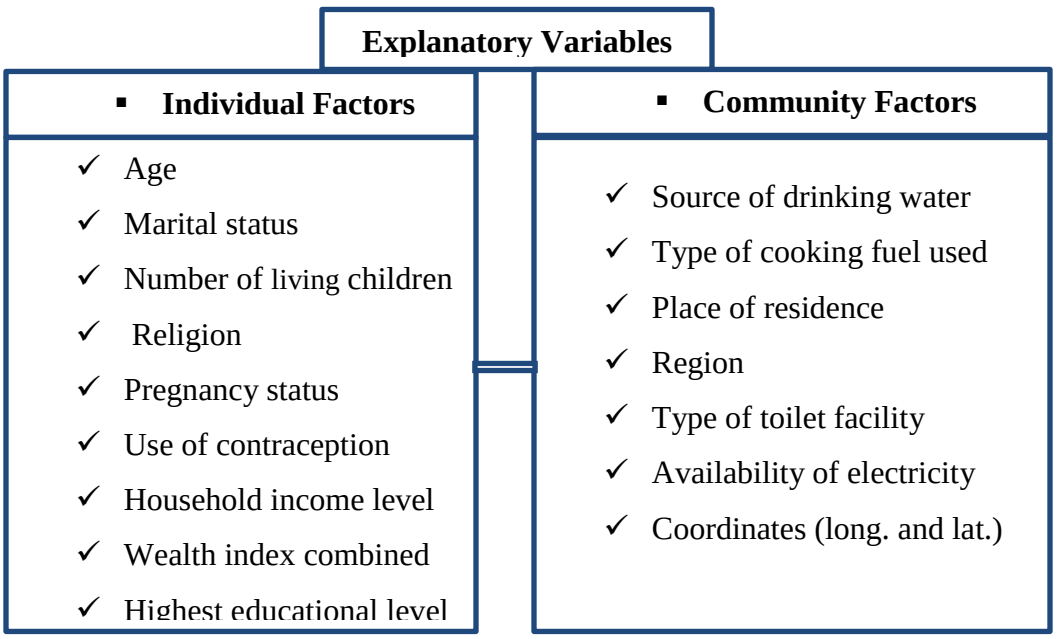
Supadi, and Wijaningsih 2019). For women of reproductive age, anaemia is commonly defined as a hemoglobin level: below 12.0 g/dL for non-pregnant women and below 11.0 g/dL for pregnant women.

Based on these standards, anaemia status in the study population can be categorized into several levels (Ejigu et al., 2018) for more detailed analysis:

i. **None anaemia:** Hemoglobin levels above the WHO threshold for anaemia.
ii. **Mild anaemia:** Hemoglobin levels slightly below the normal range.
iii. **Moderate anaemia:** Hemoglobin levels indicate a more significant decrease, which could have notable health implications.
iv. **Severe anaemia:** Hemoglobin levels that are substantially lower than normal and constitute a serious medical condition.

3.6. Explanatory Variables of the Study

These variables are chosen based on theoretical relevance and empirical evidence indicating a possible association with anaemia. The Variables coding are listed in Table 3.1 in the appendix. Here are factors of anaemia divided into Individual Factors and Community Factors based on the provided list:



3.7. Data Analysis

3.7.1. Spatial Analysis

Spatial data have different features such as point, line, area, and volume. Point: Precise locations in space indicated by a dot on a map; Line: a sequential collection of connected points like roads and rivers; Area: a region enclosed by lines like counties, states, and districts one feature of spatial data that has been considered in this study (Xu et al., 2023).

Finally, volume is a spatial data feature of an object with three dimensions, which is common for geologic formation. In spatial data analysis, locations close together in space often have similar values of outcome variables while locations far apart are often different. Everything is related to everything else, but near things are more related than far thing. This law succinctly defines the statistical notion of (positive) spatial autocorrelation, in which pairs of observations taken nearby are more alike than those taken farther apart do. This spatial correlation must be taken into account for spatial analyses. Spatial analysis is an analysis, which includes the influence of space into the analysis. All statistical methods for spatial data have to take the spatial arrangement, and the resulting correlations, of the observations into consideration to provide accurate and meaningful conclusions based on the analysis (Diggle, 2005).

3.7.2. Spatial Autocorrelation

Spatial autocorrelation measures offer additional insight into the interdependence of spatial data. These measures quantify the correlation of a spatial random field Y(s) with itself at different locations (Ejigu et al., 2018). They can be very useful for obtaining information at exact locations (point-referenced data) or measurements that characterize area-type data (real data). Different statistics have been developed to test for the presence and magnitude of spatial association among areal units. Moran's *I* is widely used, and variations of it relate to likelihood ratio tests and best invariant tests for particular models of correlation for normally distributed random variables. The value of Moran's *I* is between -1 and 1. A value close to 1 indicates strong positive spatial autocorrelation (clustered anaemia), while a value close to -1 indicates strong negative spatial autocorrelation (scattered anaemia). If Moran is *I* close to 0, it indicates that, there is no spatial autocorrelation. Having the spatial proximity matrix constructed, Moran's *I* statistics is a measure of global indexes of spatial autocorrelation with spatial proximity and similarity between areas *i* and *j* (Sawada, 2001). These include global distance-based measures

such as Moran's *I*, and Geary's C and the presence of spatial dependence is tested using Moran's I statistic

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (Y_{ij} - \bar{Y})(Y_j - \bar{Y})}{(\sum_{i}^{n} \sum_{j}^{n} w_{ij}) \sum (Y_i - \bar{Y})} \text{..(1)}$$

Where Yi represents the vector of observations at n different locations, and w_{ij} are elements of a spatial weight matrix. Values of Moran's I are assessed by a test statistic (the Moran's I standard deviate) which indicates the statistical significance of spatial autocorrelation in model residuals. In this study, Moran's I is calculated after we aggregate the number of anaemic subjects by survey cluster.

The Getis-Ord $Gi*$local statistic is an alternative to the local Moran's *I* index to determine the type of spatial cluster, i.e. hotspot or cold spot and the Geary's C is given by:

$$C = \frac{n-1}{2 \sum i \neq j} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} (Y_i - y_j)}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \text{..(2)}$$

The measure C being the ratio of the two-weighted sum of squares is never negative. It can be shown that E(C) = 1 under the assumption of no spatial association. Small value of C away from the mean 1 indicates positive spatial association. In this study areas at high risk (hotspot) of anaemia among reproductive age of women (the statistical output with high Gi∗) and areas at low risk (cold spot) of anaemia among reproductive age of women (the statistical output with low Gi∗ will be detected) (Murdock et al., 2017).

In the test of spatial significance or autocorrelation, the null hypothesis states that there is a significant association with anaemia among women of reproductive age. In contrast, the alternative hypothesis states that there is no spatial association or dependence among women of reproductive age. The research hypothesis for this study states that "there is an association or dependence between the characteristics of women of reproductive age and the prevalence of anaemia" by using Ethiopian mini demographic and health survey 2019 data. Spatial autocorrelation in the women of reproductive age is considered to be present when the test statistic such as Moran's *I* takes on a larger value, compared to what would be expected under the null hypothesis of no spatial association (Oliveira et al., 2018).

3.7.2.1. Weighted Matrix to Test Spatial Autocorrelation

Weighted matrix, sometimes called Contiguity matrix, describes the relationship between districts i and j in the specified area. The $(i,j)^{th}$element of a spatial proximity matrix W, denotedw_{ij}, quantifies the spatial dependence between regions i and j, and collectively, the w_{ij} define a neighborhood structure over the entire area. The spatial correlation parameter and W= (w_{ij}) is a neighborhood matrix for the areal units, which can be defined as:

$$\boldsymbol{w_{ij}} = \begin{cases} \mathbf{1}, & \boldsymbol{if\ regions\ share\ common\ boundary, i \neq j} \\ \mathbf{0}, & \boldsymbol{otherwise} \end{cases} \quad \text{……..(3)}$$

In this case the symmetric properties of W are established because of that w_{ji} = w_{ij} and its diagonal elements equal to zero, being the similarity of the i^{th} region with itself w_{ii} = 0.

3.7.3. Ordinal Logistic Regression (OLR) Model

The logistic regression model is a model used to study the association between a categorical dependent variable and any set of independent variables with any data type and ordinal logistic regression when a dependent variable has only two values, more than two values and more than two values with having natural order or rank respectively (Mehrolia, et al., 2021). As the dependent variable in this study, anemia status belongs to a variable that has four ordinal values, an ordinal logistic regression model was used to analyze the data. An ordinal logistic regression analysis deals with the association of a dependent variable with independent variables when the dependent variable has more than two categories having natural order or rank. The dependent variable Y is assumed to have an ordinal scale with J categories and X= (x_1, x_2, .., x_p) is the

vector of explanatory variables. Then the chances of the variable response of the j^{th} category of explanatory variable x , in particular, can be expressed by P, $P[Y = j|x] = \pi_j(x)$.When response categories are ordered, the logits can utilize the ordering that results from greater power and simple interpretation. The cumulative probability for Y is the probability that Y falls at or below a particular outcome category j and is given by:

$$\boldsymbol{P(Y \leq j) = \pi_1(x) + \ldots + \pi_j(x)} \text{ .....(5)}$$

Where j is the number of categories for the response variable Y

The cumulative logit model is as follows:

$$\boldsymbol{logit(P[Y \leq j|x]) = log\left(\frac{P[Y \leq j|x]}{1 - P[Y \leq j|x]}\right) = log\left(\frac{\pi_1(x) + \pi_2(x) + \cdots \pi_j(x)}{\pi_{j+1}(x) + \pi_{j+2}(x)}\right) = \alpha_j + X'\beta} \text{.....(6)}$$

Where: Y is the ordinal outcome variable (anaemia status)

$Y \leq j$ is the cumulative probability that Y falls in category j or below

α_j is the parameter for category j

β is the vector of coefficients for the predictor variables (explanatory variables of anaemia among reproductive age) X.

j=1, 2... J-1 and the probability $P[Y \leq j|x]$ can be estimated as:

$$\boldsymbol{P[Y \leq j|x] = \frac{\exp(\alpha_j + X'\beta)}{1 + \exp(\alpha_j + X'\beta)}} \text{.....(7)}$$

The cumulative probabilities do not use the final one $p(Y \leq j)$since it necessarily equals 1. The parameter β is a vector of regression coefficients describing the effect of the corresponding independent variable X on the log odds of response in category j or below. When this model fits well, it requires a single parameter rather than J-1 parameters to describe the effect of X. Because the model assumes that, the effect of X is identical (proportional odds) for all J-1 cumulative logits.

3.7.3.1.**Assumptions of Ordinal Logistic Regression Model**

- The dependent variable (anaemia status) is measured on an ordinal level.
- One or more of the independent variables are continuous, categorical or ordinal.
- No Multi-collinearity - i.e. when two or more independent variables are highly correlated with each other.
- Proportional Odds - i.e. that each independent variable has an identical effect at each cumulative split of the ordinal dependent variable.

The above assumptions should be tested in order as if a violation of the assumption is not correctable; we no longer be able to use ordinal regression (Saad and Yang, 2019).

3.7.3.2. Mixed Effect Ordinal Logistic Regression

Based on WHO (2017) for women ages 15 to 49 screened by hemoglobin levels who have inadequate levels and has defined and categorize anemia status as no anaemic, mild anaemic, moderate anaemic and severe anaemic based on the following cutoff values (g/dl) for women reproductive age with hemoglobin level for not anaemic more than 11.9 g/dl, for mild anaemic from 10 to 11.9 g/dl for moderate anaemic from 7 to 10 g/dl and for severe anaemic less than 7 g/dl. We consider the ordinal logistic regression of the dependent variable anaemia among reproductive age of women is ordered as:

$$Y_i = \begin{cases} \text{none anaemic, if hemoglobin level} > 11.9\ \text{g/dL} \\ \text{mild anaemic, if hemoglobin levelis (10 to 11.9)g/dL} \\ \text{moderate anaemic, if hemoglobin level is (7 to 10)g/dL} \\ \text{severe anaemic, if hemoglobin level is} < 7\ \text{g/dL} \end{cases}$$

The term mixed model refers to the use of both fixed and random effects in the same analysis. Fixed effects have levels that are of primary interest and will be used again if the experiments were repeated (Oliveira et al., 2018). Random effects have levels that are not of primary interest, but rather are thought of as a random selection from much larger set of levels. Hence, the term mixed logistic regression model revealed ordinal logistic regression model that consists both random and fixed effects factor (Seltman, 2014).

The model equation for a mixed-effects ordinal logistic regression to study the prevalence of anaemia among reproductive-age women can be formulated as follows:

$$\boldsymbol{logit\big(p(Y \leq j)\big) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + U_0 + U_1 Z_1 + \varepsilon \ldots\ldots\ldots\ldots\ldots\ldots (8)}$$

Where: $p(Y \leq j)$ is the cumulative probability of being in the j^{th} or lower category of anaemia severity. X_1, X_2 are fixed-effect covariates, Z_1 is the random effect variable (e.g., region). U_0, U_1 are random effects, and ε is the error term.

i. Fixed Effects:

Considers fixed effects for covariates that are believed to be associated with the distribution of anaemia among reproductive-age women, and also includes a list of predictor variables (explanatory variables) to be treated as fixed effects in the model. These could encompass individual-level factors (e.g., age, religion, marital status), household-level factors (e.g., educational attainment, source of drinking water), and community-level factors (e.g., place of residence).

ii. Random effect test

Describe the random effects to account for data clustering, such as the clustering of individuals within regions. These random intercepts allow for the capturing of variability at different levels of the data hierarchy. Incorporate random effects to account for clustering or unobserved heterogeneity at different levels. For example, we may include random effects for regions to capture variations in anaemia distribution that cannot be explained by the fixed effects alone. In a random effect test, the variables of interest are treated as random or varying across observations. Random effects are used to account for unobserved factors that are common across all individuals in the sample (Geisser, 1974). The models of random effect estimate the average effect of the independent variables on the dependent variable across all individuals in the sample. Also is useful when the focus is on generalizing the results to a broader population. With the random effect test, the study was focused on estimating the variation in anaemia levels among reproductive-age of women with their socioeconomic, geographic, and demographic factors, and the study used an appropriate statistical model which is a mixed effect ordinal logistic regression model to account for the hierarchical structure of the data.

3.8. Parameter Estimation

The study uses generalized linear mixed models (GLMM) techniques to estimate fixed effects, representing the impact of covariates on the ordinal status of anaemia among women of reproductive age with varying severity levels of anaemia, estimate spatial parameters that capture the spatial autocorrelation and maximum likelihood estimation (MLE).

3.8.1. Generalized Linear Mixed Model (GLMM)

A GLMM is a statistical model that accommodates both fixed effects, which are consistent and predictable factors that impact your response variable for all individuals or units (e.g., age, educational attainment), and random effects which are random factors that can vary across groups, or clusters (e.g., variation between different communities or health centers). In the case of this study, a potential effect variable could be the region where the women are located this will account for the correlation between women within the same region and help to model the variability at the group level. GLMMs are an extension of linear mixed models to allow response variables from different distributions and alternatively, we will think of GLMMs as an extension of generalized linear models (e.g., logistic regression) to include both fixed and random effects (hence mixed models) (McCulloch, 1996). The general form of the model in matrix notation is:

$$y = X\beta + Zu + \varepsilon \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

Where: y is the response variable vector or anaemia status, X is the design matrix for fixed effects. β is the vector of fixed effect coefficients, Z is the design matrix for random effects, u is the vector of random effects, and $\varepsilon$is the vector of residuals. The random components u and ε introduce the mixed aspect of the model, with u accounting for the variability between clusters (e.g., different regions) and ε representing the individual variability. In this study anaemia among women of reproductive age, a GLMM will be used to model the probability of different anaemia severity levels as an ordinal response.

3.8.2. Maximum Likelihood Estimation (MLE)

Maximum Likelihood Estimation (MLE) is a method of estimating the parameters of a statistical model (Andrade and Silva, 2017). In this study of anaemia among women of reproductive age, MLE can be used to find the parameter values that make the observed data most probable under a specified statistical model. MLE is based on the likelihood function, which measures the

probability of observing the data given certain parameter values (Zanakis and Kyparisis, 1986). The estimators are obtained by finding the parameter values that maximize this likelihood function (Luque-Fernandez et al., 2018). The underlying principle is that the observed data should be as probable as possible under the model.

In this study, we have a dataset containing measures of anaemia among women of reproductive age and associated covariates (e.g., age, educational attainment, marital-status, etc.). Let$Y_1, Y_2, \ldots \ldots, Y_n$ represent the observed anaemia status of n women of reproductive-age and let θ be the parameter or vector of parameters of the model we wish to estimate. The likelihood function $L(\theta)$for the observed data is:

$$\boldsymbol{L(\theta) = f(Y_1, Y_2, \ldots \ldots \ldots Y_n|\theta)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ldots \ldots \textbf{(10)}$$

Where, f is the joint probability density function or probability mass function of observed data given of the parameters.

3.9. Model Diagnostics and Validation

3.9.1. Goodness of Fit Test

It is useful to be able to judge whether a model is a good fit for the data. In the generalized linear mixed model (GLMM), there are no exact suggested ways to evaluate the goodness of fit of the model (Pan and Lin, 2005). The appropriate one is to check whether over or under dispersion in our data by using the goodness of fit test statistics of the Pearson Chi-square over the degree of freedom (Payne et al., 2018). In this study, the deviance value was also used to assess the goodness of fit of the model. The maximum likelihood procedure produces a statistic known as a deviance that indicates how well the model fits the data (Cole, Chu, and Greenland, 2014). We can compare models statistically using their deviances as:

$$\boldsymbol{D = D_o - D_1 = -2\log(\frac{\lambda_o}{\lambda_1})} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \textbf{(11)}$$

Where λ_o andλ_1, are the likelihoods for the null and alternative hypothesis. The difference in the deviances of two nested models has a chi-square distribution, with degrees of freedom equal to the difference in the number of parameters estimated in the two models (Signorovitch et al., 2013). This can be used to perform a formal chi-square test to check whether the more general

model fits significantly better than the simpler model (McNeish, 2020). In general, models with lower deviance values fit better than models with higher deviance. Similarly, the overall model evaluation will be also examined using Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) (Fernández-García et al., 2014).The smaller the value, the better the model will be fitted.

I. Akakie information criterion (AIC)

The AIC is a measure of fit that is used to assess models, this measure uses the log-likelihood with a penalizing term that is associated with many variables (Grueber et al., 2011). It estimates the relative amount of information lost by a given model: the less information a model loses the higher the quality of that model. Algebraically,

$$AIC = -ln(likelihood) + 2k \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (12)$$

Where k is the model degrees of freedom calculated as the rank of the variance-covariance matrix of the parameters.

Where k is the model degrees of freedom calculated as the rank of the variance-covariance matrix of the parameters.

II. Bayesian Information Criterion (BIC)

BIC is another model selection method, which uses penalty terms associated with the number of parameters, p, and the sample size, n. BIC is also known as the Schwarz Information Criterion (Wasserman, 2000). Is a criterion for a model selection among a finite set of models. it is based in part on the likelihood function and it is closely related to AIC.

$$BIC = -2\ln(likelihood) + l(N)k \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (13)$$

Where k is the model degrees of freedom calculated as the rank of the variance-covariance matrix of the parameters and N is the number of observations used in estimation.

3.9.2. Brant Test

The Brant test, also known as the Test for the Proportional Odds Assumption, is a statistical test used in the context of ordinal logistic regression to assess whether the assumption of proportional odds holds for the model. In ordinal logistic regression, the proportional odds assumption states that the effects of the independent variables on the response variable are consistent across all levels of the response variable (Liu et al., 2023). This means that the relationship between the independent variables and the odds of being in a particular category versus being in a higher category remains constant across all levels of the response variable.

The Brant test evaluates this assumption by comparing the coefficients of the independent variables in binary logistic regression models for each pair of adjacent response categories with the coefficients in the full ordinal logistic regression model. If the coefficients significantly differ between the models, it suggests a violation of the proportional odds assumption (Dolgun and Saracbasi, 2014). In essence, the Brant test is a way to check whether the assumption of proportional odds, which is a key assumption in ordinal logistic regression, is valid for the model being analyzed. If the test indicates a violation of this assumption, it may imply that the relationship between the independent variables and the response variable varies across different levels of the response variable (Lelisho, Wogi, and Tareke, 2022).

CHAPTER FOUR

4. RESULTS

This section provides, the data introduced earlier are analyzed, and the results of the analysis based on mixed effect ordinal logistic regression and spatial data analysis techniques were being presented. Remained that the main objective of the study is to assess the spatial distribution and determinant factors of anaemia among reproductive-age of women in Ethiopia using the EMDHS-2019 data.

4.1. Descriptive

When discussing the descriptive statistics results of anaemia status distribution among women of reproductive age in Ethiopia. From the total of 8885 women of reproductive age with anaemia status test results were included in this study 5090 (57.28%) were with no anaemia, 1778 (20.01%) were with mild anaemia, 534 (6.01%) were with moderate anaemia and remaining 1483(16.7%) were with severe anaemia in Ethiopia. It is essential to delve into the detailed analysis of the data presented in Table 4.1 offers insights into the percent of anaemia across anaemia status below.

Table 4.1: Distribution of anaemia status for women of reproductive age in Ethiopia from EMDHS 2019

Anaemia Status	Frequency	Percent
No Anaemia	5090	57.28%
Mild	1778	20.01%
Moderate	534	6.01%
Severe	1483	16.70%

The The descriptive statistics results of anaemia distribution among women of reproductive age in Ethiopia, in the data presented in Table 4.2 have the percentage of anaemia across Ethiopia regions and the severity levels observed. This comprehensive discussion has covered various aspects such as the percentage of anaemia across regions, severity levels, and count numbers. The table showcases a wide variation in the percentage of anaemia status among reproductive-age of women by regions in Ethiopia. The results show that an exhibit alarmingly high level of

anaemia percentage was observed in Somali, Harari, and Gambela regions recording a staggering 92.8%, 75.6%, and 15.6% respectively of severe anaemia cases among reproductive age of women. On the other hand, regions such as Addis Ababa, Tigray, and SNNPR have relatively lower percentage levels of anaemia recording 0.48%, 0.68%, and 0.69% respectively, indicating regional disparities that warrant attention.

Table 4.2: Percentage of Anaemia status by region among reproductive age of women

Regions	**Percentage of Anaemia Status by Regions**				
	No	Mild	Moderate	Severe	Total
Tigray	78.9 (578)	17.2 (126)	3.27 (24)	0.682 (5)	733
Afar	72.5 (465)	8.11 (52)	8.27 (53)	11.1 (71)	641
Amhara	83.3 (790)	12.1 (115)	3.27 (31)	1.27 (12)	948
Oromia	87.1 (917)	8.27 (86)	3.52 (37)	1.14 (12)	1052
Somali	3.75 (24)	1.88 (12)	1.56 (10)	**92.8 (594)**	640
Benishangul	4.15 (31)	80.7 (603)	7.63 (57)	7.50 (56)	747
SNNPR	87.9 (886)	7.44 (75)	3.97 (40)	0.694 (7)	1008
Gambela	7.47 (54)	69.0 (499)	7.88 (57)	**15.6 (113)**	723
Harari	5.77 (44)	5.90 (45)	12.7 (97)	**75.6 (577)**	763
Addis Ababa	78.5 (642)	18.6 (152)	2.44 (20)	0.489 (4)	818
Dire Dawa	81.3 (660)	1.48 (12)	13.3 (108)	3.94 (32)	812

The data in Table 4.2 illustrates the distribution of anaemia across four severity levels: No anaemia, mild anaemia, moderate anaemia and severe anaemia. And the table provides both count numbers and percentage breakdowns for each region and severity level. This uses to assessing the magnitude of the anaemia problem in each region. In Tigray, high percentage of no anaemia cases (78.9%), the count number for mild anaemia (126) is notable, indicating a substantial number of affected individuals.

Table 4.3: Prevalence of anaemia status by predictors in reproductive age women

Predictors	Anaemia status in women reproductive age			
	No anaemia	Mild	Moderate	Severe
Age in 5 year groups				
15-19	59.2	18.0	3.62	19.1
20-24	57.0	15.8	10.1	17.0
25-29	56.6	18.3	8.39	16.7
30-34	54.5	21.6	6.09	17.8
35-39	57.2	24.3	3.66	14.9
40-44	58.7	24.8	2.94	13.6
45-49	57.1	27.7	3.90	11.3
Type of place of residence				
Urban	56.1	14.7	10.7	18.4
Rural	57.9	22.6	3.66	15.8
Religion				
Orthodox	65.1	21.3	6.58	6.99
Catholic	47.4	24.4	10.3	17.9
Protestant	62.0	24.5	7.07	6.37
Muslim	48.0	16.2	4.95	30.8
Traditional	61.7	38.3	-	-
Other	44.4	25.9	11.1	18.5
Educational attainment				
No education	54.5	23.6	1.40	20.5
Incomplete primary	55.4	32.0	0.881	11.8
Complete primary	99.4	-	0.197	0.394
Incomplete secondary	65.3	-	0.0981	34.6
Complete secondary	58.5	1.54	37.7	2.31
Higher	38.5	1.46	54.2	5.86

Wealth index combined				
Poorest	50.2	20.6	2.51	26.7
Poorer	60.3	26.0	1.72	12.1
Middle	62.5	25.1	3.39	8.99
Richer	59.5	23.5	4.76	12.2
Richest	57.5	13.0	12.2	17.3
Source of drinking water				
Protected	56.9	20.9	7.17	15.0
Un protected	58.2	17.8	3.12	20.9
Pregnancy status				
No	57.7	20.1	6.09	16.0
Yes	52.8	18.6	5.23	23.3
Use of contraception				
No	56.0	19.2	5.81	19.0
Yes	61.5	22.6	6.66	9.26
Marital status				
Never in union	59.8	15	7.70	17.5
Married	57.0	21.0	5.38	16.6
Divorced	54.2	24.2	6.82	14.8
Widowed	48.5	33.9	2.64	15.0
Household has radio				
No	56.4	20.7	5.32	17.6
Yes	59.4	18.3	7.63	14.7
Household has Tv				
No	57.2	22.9	3.76	16.1
Yes	57.6	12.0	12.2	18.2
Household has refrigerator				
No	56.8	21.8	4.38	17.1
Yes	60	10.4	14.8	14.7

Table 4.3 provides a summary of the percentage of anaemia by anaemia status level under different categorical covariates including age in 5-year groups, type of place of residence, religion, educational attainment, wealth index combined, pregnancy status, use of contraception, marital status. Based on the ordinal level of anaemia status the higher prevalence of anaemia was observed in type of place of residence (urban) areas compared with individuals who live in rural areas. Furthermore, anaemia prevalence decreases as the wealth index increases and the anaemia prevalence is increase when the educational level increases for women of reproductive age. For severe anaemia status, the prevalence of anaemia is higher who were active in pregnancy status and for women who use contraception. To see whether this difference is statistically significant or not, Bi-variable or multivariable mixed effect ordinal logistic regression model (Tables 4.5) and a generalized linear mixed-effect model (Tables 4.7) with including all potential predictors simultaneously have been fitted. The results are represented in Tables 4.5 and Table 4.7.

4.2. Spatial Analysis

4.2.1. Spatial Distribution of Anaemia

To get a general insight into the spatial clustering of anaemia, a global spatial statistic was estimated using Moran's *I* statistic (equation 1). Figure 4.1 below presents global spatial autocorrelation or Moran's *I* for the distribution of anaemia among reproductive age of women. The spatial distribution of anaemia among reproductive age of women in Ethiopia was identified to be clustered (Global Moran's I = 0.246652, p-value < 0.001). The outputs were automatically produced keys on the right and left sides of each panel. Given the z-score of 2.8, it is clear that there is less than 1% likelihood that this clustered pattern could be the result of chance. The bright red and blue colors to the end tails indicate an increased significance level. The Moran's Index value of 0.246652 indicates that positive spatial autocorrelation. This means that there is a tendency for similar values to be located near each other more than would be expected under a random spatial distribution. Although the value shows a moderate level of autocorrelation, it signifies that the spatial pattern of the analyzed variable is not random but exhibits some degree of clustering. The z score measures how many standard deviations the observed Moran's Index is away from the expected Moran's Index under the null hypothesis of spatial randomness. A z score of 2.8 suggests that the observed spatial autocorrelation is 2.8 standard deviations away from the expected value if there were no spatial autocorrelation. Typically, a z score above

approximately 1.96 or below -1.96 indicates statistical significance at the 5% level. However, this z score corresponds to a significance level (p-value) below 0.01, denoting high statistical significance and that the observed clustering is unlikely to have occurred by chance or random. The p-value provides the probability of observing the given level of spatial autocorrelation (or more extreme) if the null hypothesis of no spatial autocorrelation were true. A p-value of 0.005519 indicates that there is a very low probability that the observed spatial autocorrelation pattern occurred due to random chance. This statistically significant result (since the p-value is less than the typical alpha level of 0.05) strongly suggests that the observed clustering pattern is not a product of randomness.

The statement that "there is a less than 1% likelihood that this clustered pattern could be the result of random chance" aligns with the reported p-value of 0.005519. The p-value indicates that there is approximately a 0.55% chance (which is less than 1%) that the observed degree of clustering could occur if the spatial distribution were random. Therefore, the spatial autocorrelation analysis provides strong evidence of significant clustering, and the spatial pattern observed in the data is unlikely to be due to randomness. This reinforces the conclusion that the variable being analyzed exhibits a significant spatially clustered pattern across the regions.

4.2.2. Getis-Ord Gi∗ Statistical Analysis of Anaemia

The global Morans' I statistic informs whether there is a spatial aggregation of anaemia among reproductive-age women, it does not allow for the identification of regions classification according to the level of their significance. Therefore, to identify high clusters (the hot-spots) and low clusters (the cold-spots) regions for the distribution of anaemia among reproductive age of women, the study applied the Getis-Ord Gi∗ statistic. A map showing the distribution of spatial clusters of anaemia among reproductive-age women in Ethiopia is presented in (Fig. 4.2). In this figure we see Gi_Bin at the right-top of the map shows that it is the type of local spatial autocorrelation statistic used in spatial statistics and geographic information science to identify hot spots, cold spots, and spatial outliers in data. This statistic is often employed in hot spot analysis to detect significant clusters of high values (hot spots) and low values (cold spots) in a spatial data set.

Based on the Getis-Ord Gi* statistical analysis, this study identified hotspots and cold spot areas of anaemia among reproductive-age women in Ethiopia. Accordingly, the red colors indicate the

significant hotspot area (higher cluster of anaemia), which was found in the Somalia and Harari regions. In contrast, the blue color indicates significant cold spot areas (low cluster of anaemia), located in the Amhara, Tigray, Oromia, SNNPR, Benishangulgumuz, and Afar regions (Figure 4.2).

Figure 4.3 presents graphical plots of the Moran's I test statistic values as a function of distance (in meters). As presented in Figure 4.3, significant local clustering of anaemia distribution occurs between pairs of clusters within 1-200 km distance lags; the highest mean local Moran's *I* values are observed at a 250 km distance lag.

Figure 4.4 below presents the observed anaemia distribution by the 2019 EMDHS survey clusters. The result shows that the distribution of anaemia among reproductive age of women varies from cluster to cluster through regions. This result also shows a general insight into the distribution of anaemia among reproductive-age women by regions of the country and presents the distribution of anaemia by regions of the country. The map reveals that the eastern part of the country had higher anaemia distribution among women of reproductive age than the other direction of the country. The red color in Figure 4.4 identifies the Somali region with hotspot anaemia distribution among women of reproductive age. Also, the Harari region, which is represented with the bold pink color, is higher in anaemia spatial distribution among women of reproductive age.

Given the z score of 2.7750859719, there is less than a 1% likelihood that this clustered pattern could be the result of random chance. .

Table 4.4: Results of Global Moran's I Statistics

Global Moran's I Summary	
Moran's Index	0.246652
Z-Score	2.7750859719
P-value	0.00000

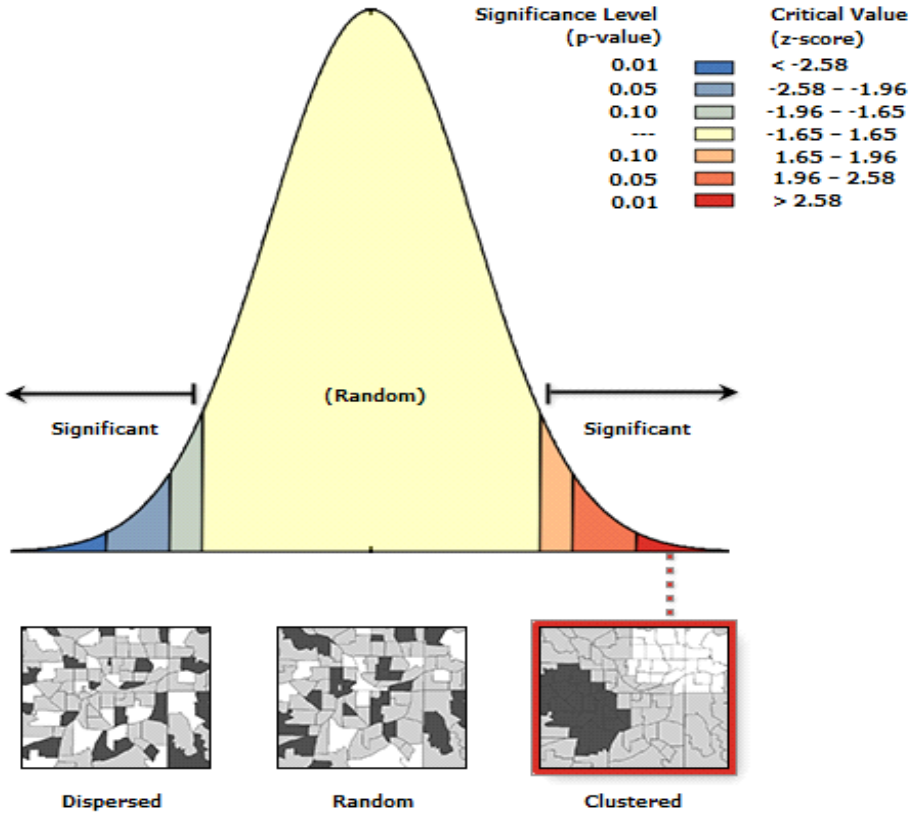


Figure 4.1: Spatial autocorrelation analysis of anaemia among reproductive age of women in Ethiopia using EMDHS 2019 data

4.2.3. Testing for Spatial Autocorrelation

The Moran's I being among the most widely implemented measures of spatial autocorrelation between neighboring districts as briefly discussed in the methodology part in section (3.7.2), was used. In this section, our focus is on their application to particular data analysis, the essential task being to seek for spatial distribution.

Spatial autocorrelation analysis includes tests and visualization of global test (Moran's I and Geary's C) statistics. First, the global Moran's I test statistics were computed to test the null hypothesis ($H_0: \rho = 0$) of no significant clustering of anaemia incidence in the entire study region ($\alpha = 0.05$). The test was repeated using a diagnostic for spatial dependence to validate the consistency of the results.

The main objective of estimating spatial autocorrelation coefficients (global and local) is to measure the degree of spatial autocorrelation between anaemia incidence rates between

neighboring areas, to look for spatial patterns or to diagnose spatial dependence in the regression model. The tests are performed under the assumption of normality and the null hypothesis stipulating spatial independence (uncorrelated error terms) for the data under consideration. The estimated results of Moran's I are also used to identify the model.

Moran’s I suggests the presence of statistically significant positive spatial autocorrelation in the distribution of anaemia among women of reproductive age in Ethiopia. This implies that there are regional patterns or clusters where anaemia prevalence is higher or lower than expected by random distribution. This finding could reflect underlying factors such as socio-economic, geographical, or healthcare access that vary spatially and influence anaemia rates.

The Moran’s I scatter plot shows a weak but statistically significant positive spatial autocorrelation for anaemia status. This suggests some clustering in the distribution of anaemia rates across the regions, meaning neighboring regions tend to have similar anaemia statuses, but the spatial dependence is not very strong. Further exploration using local spatial statistics could help identify specific regions where clustering is more pronounced.

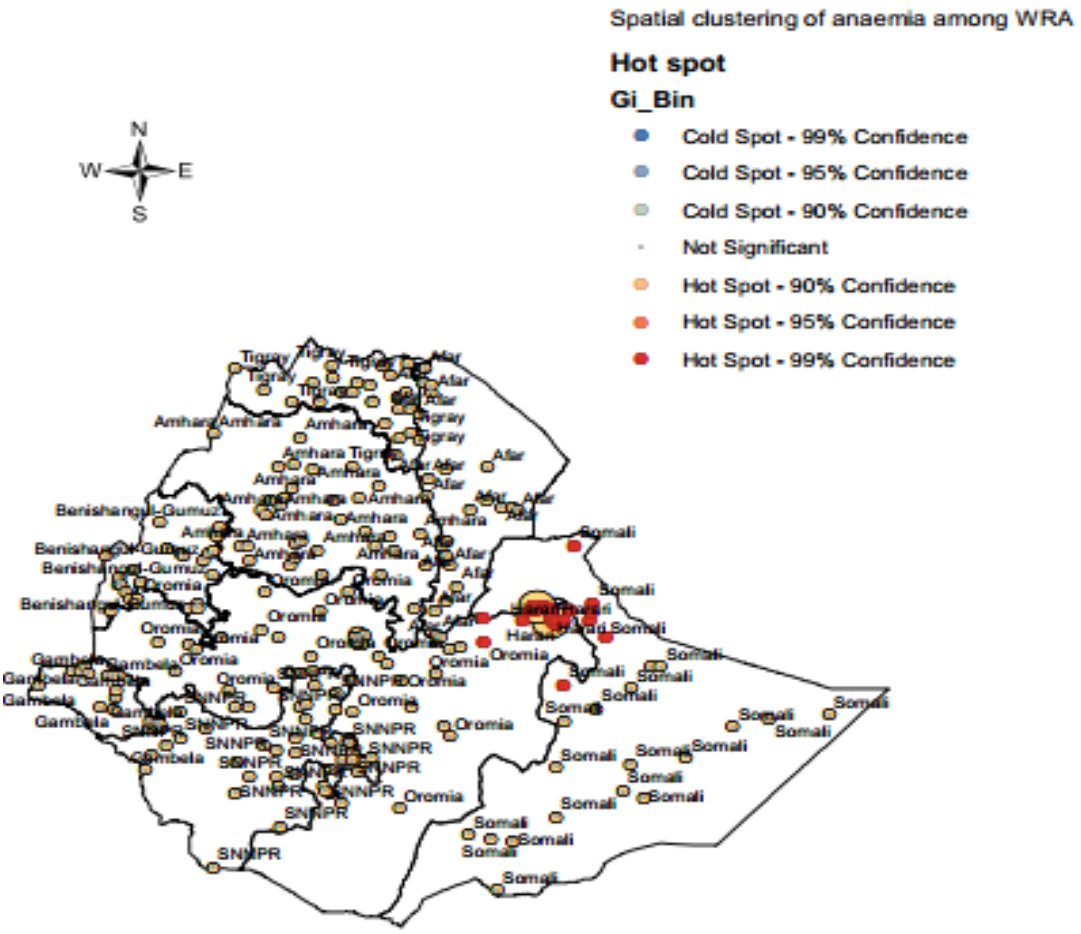


Figure 4.2: Hotspot and Cold spot areas anaemia across regions among reproductive age of women in Ethiopia using EMDHS 2019 data.

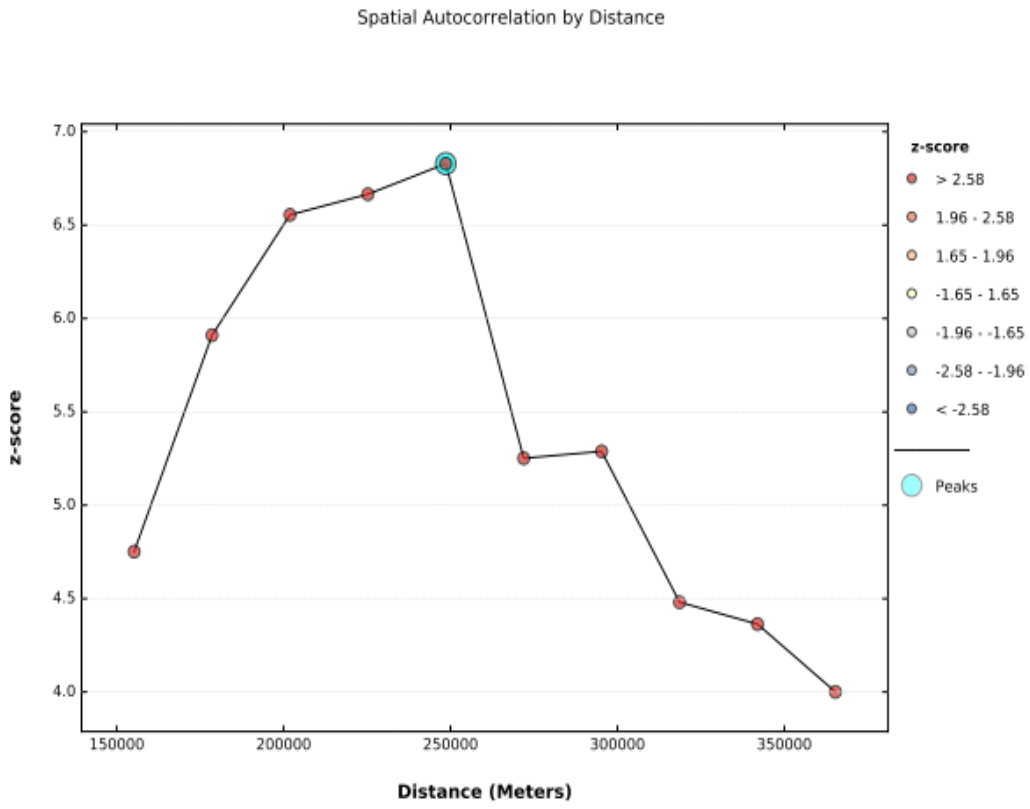


Figure 4.3: spatial autocorrelation as a function of distance

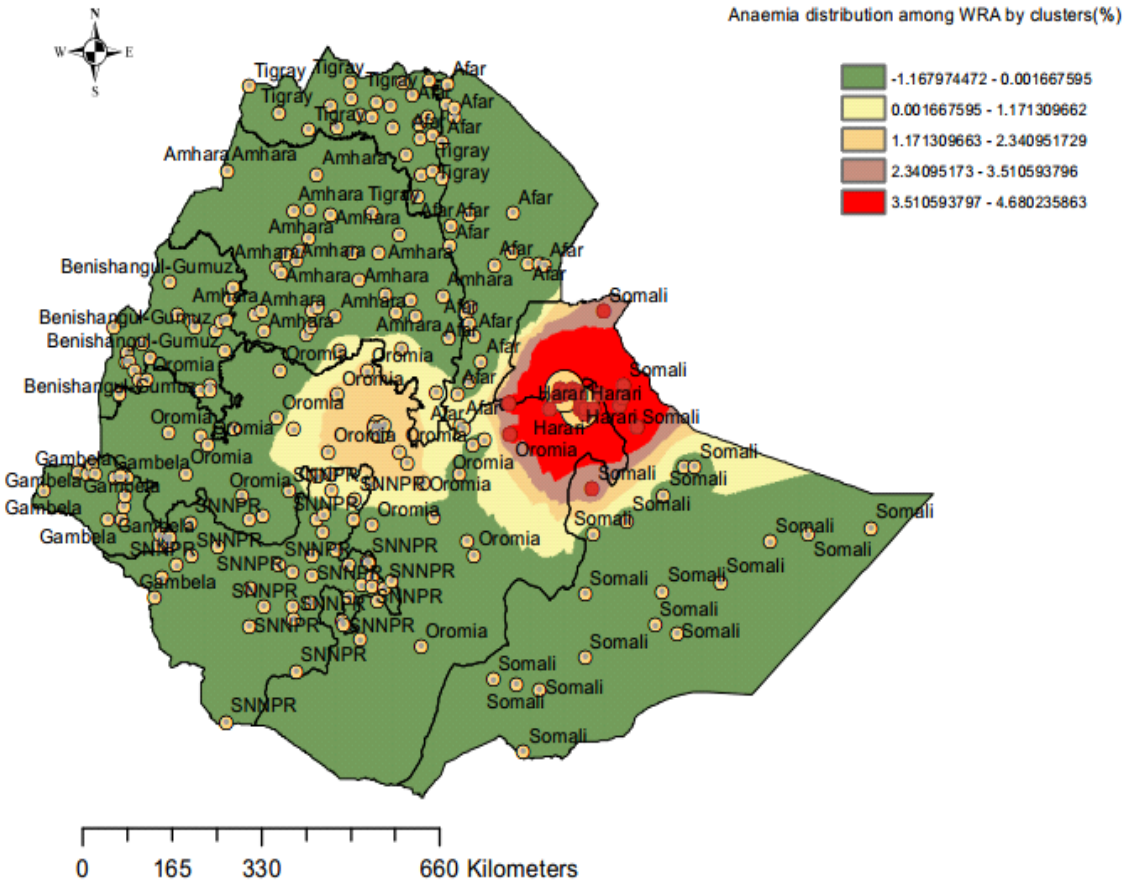


Figure 4.4: Anaemia distribution among reproductive age of women for the 2019 EMDHS

4.3. Ordinal Logistic Regression Model

The Ordinal Logistic Regression (OLR) model, also known as the proportional odds model, is a widely used method for analyzing outcomes that have a natural order but no fixed interval between categories. In this study, the OLR model was employed to investigate the relationship between multiple predictor variables and anemia status among women reproductive age, which is categorized into four ordinal levels. This model assumes that the relationship between the independent variables and the log odds of being in a higher category is constant across all thresholds of the outcome variable. The OLR model is particularly suitable for ordinal outcomes because it preserves the inherent ordering of the response categories while allowing for the inclusion of both continuous and categorical predictors. This model helps to quantify the effect of predictor variables such as age, region, wealth index and household characteristics on the

likelihood of severe anemia severity, providing insights into the factors associated with anemia in the study population.

Table 4.5: Bi-variable and multivariable mixed-effect ordinal logistic regression model of anaemia among reproductive-age women in Ethiopia

Variables	Crude odds ratio			Adjusted odds ratio		
	OR.	95% CI		OR.	95% CI	p-value
Region						
Tigray	Ref	Ref .		Ref	Ref	Ref
Afar	.425	[0.333, 0.542]		1.057	[0..79, 1.416]	.709
Amhara	.462	[0.369, 0.578]		**.683**	**[0.539,.866]**	**.002**
Oromia	.476	[0.382, 0.592]		.791	[0.615, 1.016]	.067
Somali	.385	[0.301, 0.493]		1.026	[0.764, 1.38]	.863
Benishangul	.577	[0.457, 0.727]		1.022	[0.79, 1.323]	.866
SNNPR	.511	[0.41, 0.635]		**.743**	**[0.568, 0.971]**	**.03**
Gambela	.802	[0.636, 1.012]		1.081	[0.826, 1.415]	.57
Harari	1.182	[0.95, 1.471]		.992	[0.771, 1.277]	.952
AddisAbaba	2.272	[1.849, 2.79]		**.707**	**[0.553, 0.904]**	**.006**
Dire Dawa	1.223	[0.985, 1.517]		.976	[0.759, 1.256]	.852
Age in 5-yr groups						
15-19	Ref	Ref		Ref	Ref	Ref
20-24	1.509	[1.306, 1.743]		**1.648**	**[1.401, 1.939]**	**0**
25-29	1.169	[1.015, 1.347]		**1.678**	**[1.403, 2.007]**	**0**
30-34	.818	[0.695, 0.962]		**1.392**	**[1.127, 1.719]**	**.002**
35-39	.684	[0.577, 0.811]		1.256	[0.999, 1.578]	.051
40-44	.645	[0.531, 0.783]		1.272	[0.981, 1.648]	.069
45-49	.636	[0.512, 0.789]		1.305	[0.984, 1.731]	.065
Type of place of residence						
Rural	Ref	Ref		Ref	Ref	Ref
Urban	3.734	[3.377, 4.129]		**1.22**	**[1.025, 1.452]**	**.025**
Religion						
Orthodox	Ref	Ref		Ref	Ref	Ref
Catholic	1.462	[0.933, 2.291]		**1.997**	**[1.232, 3.235]**	**.005**
Protestant	.74	[0.651, 0.841]		1.105	[0.936, 1.305]	.237
Muslim	.454	[0.408, 0.505]		**.604**	**[0.522, 0.699]**	**0**
Traditional	.292	[0.163, 0.522]		.578	[0.306, 1.09]	.09
Other	1.803	[0.876, 3.711]		**3.069**	**[1.419, 6.637]**	**.004**
Type of toilet facility						
No facility	Ref	Ref .		Ref	Ref	Ref
Toilet with flush	6.01	[4.986, 7.245]		1.148	[0.898, 1.468]	.271

Fit latrine	1.798	[1.618, 1.997]		.971	[0.843, 1.119]	.686
Wealth index combined						
Poorest	Ref	Ref .		Ref	Ref	Ref
Poorer	1.093	[0.923, 1.294]		1.029	[0.851, 1.244]	.768
Middle	1.365	[1.149, 1.622]		**1.274**	**[1.041, 1.56]**	**.019**
Richer	1.695	[1.43, 2.009]		**1.579**	**[1.27, 1.963]**	**0**
Richest	5.407	[4.718, 6.196]		**2.409**	**[1.798, 3.228]**	**0**
Source of drinking water						
Protected	Ref	Ref .		Ref	Ref	Ref
Unprotected	.493	[0.442, 0.549]		.923	[0.813, 1.048]	.217
Type of cooking fuel						
Clean	Ref	Ref .		Ref	Ref	Ref
Solid	.198	[0.177, 0.222]		**.541**	**[0.458, 0.639]**	**0**
Pregnancy status						
No	Ref	Ref .		Ref	Ref	Ref
Yes	.846	[0.718, 0.998]		1.026	[0.852, 1.236]	.787
Number of living child	.823	[0.807, 0.84]		**.923**	**[0.894, 0.954]**	**0**
Use of contraception						
No	Ref	Ref .		Ref	Ref	Ref
Yes	1.039	[0.933, 1.158]		1.022	[.896, 1.165]	.747
Marital status						
Never in union	Ref	Ref .		Ref	Ref	Ref
Married	.445	[0.4, .495]		**.636**	**[0.541, 0.747]**	**0**
Divorced	.512	[0.42, .625]		**.501**	**[0.401, 0.625]**	**0**
Widowed	.381	[0.281, .518]		**.548**	**[0.388, 0.773]**	**.001**
Availability of electricity						
No	Ref	Ref		Ref	Ref	Ref
Yes	3.232	[2.928, 3.568]		**.799**	**[0.662, 0.964]**	**.019**
Household has radio						
No	Ref	Ref		Ref	Ref	Ref
Yes	1.604	[1.452, 1.771]		1.032	[0.923, 1.154]	.579
Household has Tv						
No	Ref	Ref .		Ref	Ref	Ref
Yes	4.039	[3.643, 4.478]		1.133	[0.921, 1.394]	.237
Household has refrigerator						
No	Ref	Ref		Ref	Ref	Ref
Yes	4.497	[4.01, 5.043]		**1.317**	**[1.105, 1.568]**	**.002**

Based on a multivariable mixed-effects ordinal logistic regression model, marital status, age, place of residence, region, wealth index, religion, and pregnancy status were significantly

associated with the level of anaemia among women of reproductive age in Ethiopia at p-value 0.05 (Table 4.5).

Individuals from the Amhara, SNNPR, and Addis Ababa regions have significantly lower odds of being in severe anaemia compared to those from Tigray (the reference group). Specifically, individuals from Amhara have 31.7% lower odds of severe anaemia, SNNPR residents have 25.7% lower odds, and those in Addis Ababa have 29.3% lower odds. These results indicate that individuals from these regions are less likely to experience severe anaemia compared to those from Tigray.

Age in five-year groups 20-24, 25-29, and 30-34 have significantly higher odds of being in severe anaemia compared to the youngest group, 15-19. Those aged 20-24 have 64.8% higher odds, individuals aged 25-29 have 67.8% higher odds, and those aged 30-34 have 39.2% higher odds of being in a more severe anaemia category. This suggests that anaemia severity tends to increase with age, particularly for individuals in their twenties and early thirties age..

Individuals living in urban areas have 22% higher odds of being in a higher anaemia category compared to those living in rural areas. This finding suggests that urban type of residents may be at greater risk of more severe anaemia than their rural counterparts, possibly due to differences in lifestyle, access to nutrition, or healthcare services.

Religion significantly influences anaemia status. Catholics have almost twice the odds (99.7% higher) of being in a severe anaemia category compared to Orthodox Christians (the reference group). In contrast, Muslims have 39.6% lower odds of severe anaemia compared to Orthodox Christians. Additionally, individuals practicing “other” religions have 206.9% higher odds of being in a severe anaemia category. These results highlight that religion is a crucial factor, possibly reflecting different health practices, dietary patterns, or sociocultural influences on anaemia among women of reproductive age.

The wealth index is a significant predictor of anaemia severity. Individuals in the middle wealth group have 27.4% higher odds of being in a severe anaemia category compared to the poorest individuals (reference group). The odds increase further for the richer groups who have 57.9% higher odds, and for the richest who have 140.9% higher odds. This suggests that wealthier

individuals may experience more severe anaemia, which could be attributed to varying lifestyle factors or differences in access to healthcare and nutritional habits.

Those using solid cooking fuels (e.g., wood, charcoal) have 45.9% lower odds of being in a severe anaemia category compared to those using clean fuels (e.g., gas, electricity). This inverse relationship indicates that households relying on solid fuels may have lower anaemia risk, potentially due to cultural or socio-economic factors associated with diet and living conditions among women of reproductive age.

For each additional living child, the odds of being in a severe anaemia category decrease by 7.7%. This suggests that families with more children may experience lower anaemia severity, possibly reflecting better family health practices, nutritional awareness, or support systems that mitigate the risk of severe anaemia.

Marital status has a strong influence on anaemia severity. Married individuals have 36.4% lower odds of being in a higher anaemia category compared to those who have never been in a union (reference group). Additionally, divorced individuals have 49.9% lower odds, and widowed individuals have 45.2% lower odds of being in a severe anaemia category. This suggests that being in or having been in a marital relationship is associated with lower anaemia risk, which may be due to better social or economic support.

Individuals with electricity access have 20.1% lower odds of being in a severe anaemia category compared to those without access to electricity. This may reflect the broader benefits of improved living conditions and better access to health-promoting resources (e.g., refrigeration of food, better cooking facilities) associated with electricity availability.

Owning a refrigerator is associated with 31.7% higher odds of being in a severe anaemia category. While this seems counterintuitive, it may indicate that wealthier households with refrigerators may also have lifestyle factors, such as a diet high in processed foods or urban living conditions that contribute to more severe anaemia despite their access to modern amenities among women of reproductive age.

4.3.1. Test for the Proportional Odds Assumption

The Brant test evaluates whether the proportional odds assumption holds in an ordinal logistic regression model. When this assumption is violated, it means the relationship between the predictor and the outcome isn't consistent across categories of the outcome variable. A significant test statistic provides evidence that the parallel regression assumption has been violated.

Table 4.6: Brant test of parallel regression assumption

Variables	chi^2	$p>chi^2$
All	2354.430	0.000
Region		
Afar	34.460	0.000
Amhara	6.590	0.037
Oromia	8.010	0.018
Somali	7.640	0.022
Benishangul	19.630	0.000
SNNPR	9.340	0.009
Gambela	13.360	0.001
Harari	32.000	0.000
AddisAbaba	29.720	0.000
Dire Dawa	23.590	0.000
Age in five year groups		
20-24	184.900	0.000
25-29	161.940	0.000
30-34	110.800	0.000
35-39	75.870	0.000
40-44	54.080	0.000
45-49	46.650	0.000
Type of place of residence		
Urban	2.860	0.239

Type of toilet facility		
Toilet with flush	4.730	0.094
Fit latrine	6.280	0.043
Wealth index combined		
Poorer	14.330	0.001
Middle	56.210	0.000
Richer	101.110	0.000
Richest	109.540	0.000
Source of drinking water		
Unprotected	2.470	0.291
Type of cooking fuel		
Solid	47.460	0.000
Pregnancy status		
Yes	0.830	0.659
Number of living children	206.740	0.000
Use of contraception		
Yes	6.100	0.047
Marital status		
Married	7.760	0.021
Divorced	18.230	0.000
Widowed	5.770	0.056
Availability of electricity		
Yes	0.450	0.799
House hold has radio		
Yes	4.270	0.118
House hold has Tv		
Yes	10.740	0.005
House hold has refrigerator		
Yes	33.350	0.000

The Brant test indicates that the proportional odds assumption is violated for several variables in the model. Specifically, for region, significant violations are observed for Afar, Amhara, Oromia, Somali, Benishangul, SNNPR, Gambela, Harari, Addis Ababa, and Dire Dawa, suggesting that the effect of region varies across outcome categories. All age groups (20-24 to 45-49) also show strong violations, indicating that age impacts the outcome differently across categories. The type of residence (urban) and source of drinking water (unprotected) do not violate the assumption, meaning their effects are consistent across categories. However, type of cooking fuel (solid), wealth index (middle, richer, richest), number of living children, use of contraception, and marital status (married, divorced) show significant violations, indicating non-proportional effects. Household characteristics like refrigerator and television ownership also violate the assumption, while pregnancy status, availability of electricity, and radio ownership do not. Given these results, adjusting the model by using alternatives like generalized ordinal logistic regression or multinomial logistic regression is recommended for variables that violate the assumption.

4.4. Generalized Linear Mixed Model (GLMM)

The Generalized Linear Mixed Model (GLMM) is an advanced statistical approach used to analyze complex data where outcomes may not follow a normal distribution and hierarchical or clustered structures are present. GLMM extends the Generalized Linear Model (GLM) by incorporating both fixed and random effects, making it suitable for handling data with random variations across different groups or clusters. This model is particularly useful when dealing with data that exhibits non-independence, such as repeated measures or multi-level data, where observations are nested within larger units (e.g., regions, individuals, or time periods). By combining the flexibility of GLMs in modeling various types of response variables (binary, categorical, count, etc.) with the ability to account for random effects, GLMM provides a robust framework for analyzing data that captures both population-level patterns and individual-level variability, ensuring more accurate and generalizable results in a wide range of research fields.

Table 4.7: Generalized linear mixed-effect model of anemia status among reproductive age of women in Ethiopia

Anaemia status	Coefficient	Std. err	Z	P>|z|	[95% conf. interval]	
Age in 5year groups						
2	.499938	.0828349	6.04	0.000	.3375846	.6622914
3	.5110133	.0913009	5.60	0.000	.3320668	.6899597
4	.3230739	.1076253	3.00	0.003	.1121323	.5340156
5	.2180075	.116511	1.87	0.061	-.0103499	.4463649
6	.2293134	.1322033	1.73	0.083	-.0298004	.4884271
7	.2506332	.1439454	1.74	0.082	-.0314946	.5327611
Type of place of residence						
	.205031	.088133	2.33	0.020	.0322934	.3777685
Religion						
2	.7107044	.2442686	2.91	0.004	.23194671	1.189462
3	.1044179	.0794763	1.31	0.189	-.0513528	.2601885
4	-.4634591	.0701223	-6.61	0.000	-.6008963	-.326022
5	-.5515398	.3224042	-1.71	0.087	-1.18344	.0803607
6	1.1307	.3924543	2.88	0.004	.3615035	1.899896
Type of toile facility						
2	.1046022	.1247238	0.84	0.402	-.139852	.3490564
3	-.0461775	.0711808	-0.65	0.517	-.1856893	.0933342
Wealth index combined						
2	.0007165	.0955584	0.01	0.994	-.1865744	.1880074
3	.2136284	.1023026	2.09	0.037	.013119	.4141378
4	.4291086	.1102027	3.89	0.000	.2131153	.645102
5	.8591041	.1481702	5.80	0.000	.568696	1.149512
Source of drinking water						
	-.0832925	.0646406	-1.29	0.198	-.2099857	.0434007
Type of cooking fuel						
	-.5983471	*.08398*	*-7.12*	*0.000*	*-.7629449*	*-.4337493*
Pregnancy status						
	.0225564	.0949293	0.24	0.812	-.1635016	.2086145
Number of living children						
	-.0793552	.0164873	-4.81	0.000	-.1116697	-.0470407
Use of contraception						
	.0135541	.0668866	0.20	0.839	-.1175412	.1446495
Marital status						
2	.4416944	.0823548	-5.36	0.000	-.6031068	-.2802821
3	-.6756758	.1134878	-5.95	0.000	-.8981077	-.4532439
4	-.5816616	.1753367	-3.32	0.001	-.9253153	-.2380079
Availability of electricity						
	-2168265	.0956523	-2.27	0.023	-.4043016	-.0293514
Household has radio						

	.0284235	.0567623	0.50	0.617	-.0828286	.1396757
Household has Tv						
	.1243206	.1048779	1.19	0.236	-.0812364	.3298776
Household has refrigator						
	.2877088	.0890919	3.23	0.001	.1130919	.4623258
/cut1	-3.654924	.1392851			-3.927918	-3.38193
/cut2	.9844308	.1333864			.7229983	1.245863
/cut3	1.869218	.134782			1.60505	2.133386
Region						
Var(_cons)	.0180146	.0110347			.0054229	.059844
LR test vs. ologit model:	***chibar2(01)= 13.17***			***Prob >= chibar2 = 0.0001***		

The results of the Generalized Linear Mixed Model (GLMM) for anaemia status indicate several significant predictors of this ordinal outcome.

The odds of having a higher anaemia status increase with age so that compared to the reference group (likely 15-19 years), individuals in the age groups 20-24, 25-29, and 30-34 show significantly higher odds of being in a higher anaemia category. The coefficients for these age groups are positive and statistically significant (e.g., 0.4999 for age group 20-24).

Living in an urban area is associated with increased odds of having a higher anaemia status compared to living in a rural area. The coefficient for urban residence is 0.2050 and is statistically significant (p=0.020), indicating that people in urban areas may have higher odds of experiencing anaemia. This could point to urban-specific factors contributing to anaemia, possibly related to nutrition or access to healthcare.

The effect of religion on anaemia status varies across categories incase individuals in Religion category 2 (possibly Muslim) and category 6 (possibly Traditional) have significantly higher odds of being in a higher anaemia category compared to the reference group, with coefficients of 0.7107 and 1.1307, respectively. In contrast, individuals in Religion category 4 (possibly Orthodox) have significantly lower odds of being in a higher anaemia category (coefficient = -0.4635, p = 0.000), suggesting that religious practices or socio-cultural factors may influence anaemia risk differently.

Wealth Index is an important protective factor against severe anaemia. As wealth increases, the odds of being in a higher anaemia category decrease. The wealthiest category (Richest) has the

largest positive coefficient (0.8591, p = 0.000), indicating a strong protective effect. The Middle and Richer wealth categories also show significant protective effects, with positive coefficients of 0.2136 and 0.4291, respectively. This suggests that higher socioeconomic status is linked to better health outcomes and lower anaemia risk.

The use of solid cooking fuel is associated with significantly lower odds of being in a higher anaemia category. The coefficient for this variable is -0.5983, which is highly significant (p = 0.000), suggesting that households using solid fuel for cooking are more likely to have severe anaemia. This could be due to the negative health effects of solid fuel use, such as exposure to indoor air pollution.

Having more living children is associated with lower odds of severe anaemia. The coefficient for this variable is -0.0794, which is statistically significant (p = 0.000). This could suggest that women with more children are more likely to be involved in activities or behaviors that reduce anaemia risk, or that household resources might be distributed in ways that improve anaemia outcomes in larger families.

Marital status has a significant impact on anaemia status. Being married, divorced, or widowed is associated with lower odds of severe anaemia compared to the reference group (possibly single or never married). The coefficients for these categories are all negative and statistically significant (e.g., -0.4417 for married individuals), indicating that marital status plays a protective role against severe anaemia.

Having access to electricity is associated with lower odds of being in a higher anaemia category. The coefficient for this variable is -0.2168 and is statistically significant (p = 0.023), indicating that households with electricity may have better living conditions, which contribute to better anaemia outcomes.

Among household assets, having a refrigerator is significantly associated with lower odds of severe anaemia, with a positive coefficient of 0.2877 (p = 0.001). In contrast, having a radio or television does not show a statistically significant effect on anaemia status. This suggests that ownership of a refrigerator, which could be linked to better food storage and nutrition, plays a role in protecting against anaemia.

The model identifies several socio-demographic and household factors that significantly affect anaemia status, with age group, religion, wealth index, cooking fuel type, number of children, marital status, and access to electricity and refrigerators all playing important roles. The mixed-effects approach accounts for regional variability, suggesting that anaemia outcomes are also influenced by unobserved regional factors.

4.5. Model Diagnostics and Validation

4.5.1. Goodness of Fit Test

The goodness of fit test is an essential step in evaluating how well a statistical model explains the observed data. For Generalized Linear Mixed Models (GLMMs), goodness of fit tests assess whether the model appropriately captures the patterns in the data or if there is a significant difference between the predicted and observed outcomes. Several methods can be used to assess the goodness of fit in GLMMs, including the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), both of which compare different models to determine the best-fitting one. Lower values of AIC and BIC indicate a better model fit.

Table 4.8: AIC and BIC model comparison

Model	**AIC**	**BIC**
OLR	14151.88	14456.84
GLMM	14159.99	14401.12

In the comparison of the Ordinal Logistic Regression (OLR) and Generalized Linear Mixed Model (GLMM), model selection was guided by penalizing complexity using the Bayesian Information Criterion (BIC), which applies a stronger penalty for the number of parameters as compared to the Akaike Information Criterion (AIC). While both models exhibited similar fit based on the AIC values, the BIC favored the GLMM, with a BIC of 14401.12 compared to 14456.84 for the OLR model. This result indicates that the GLMM strikes a better balance between model fit and complexity, making it the preferred model in this context for its ability to account for random effects while maintaining a parsimonious structure.

CHAPTER FIVE

5. DISCUSSION

This study examined factors associated with anaemia status among women of reproductive age, uncovering complex relationships between anaemia and various socio-demographic, lifestyle, and regional characteristics. These findings align with and add to previous research, offering important insights for public health interventions.

The study found a positive association between age and increased likelihood of severe anaemia, particularly among women aged 20-34. Similar age related trends in anaemia have been reported in studies from various regions in Ethiopia and other countries, where higher anaemia prevalence was observed among young adults compared to older adults (Gebremedhin and Enquselassie, 2011). This may reflect biological and nutritional vulnerabilities that are more pronounced in younger women, especially those of reproductive age, as well as the demands of pregnancy and childbirth (Saini, Singh, and Agrawal, 2019).

Pregnant women showed a higher likelihood of severe anaemia compared to non-pregnant women, which aligns with studies from the Gambia (Shitu and Terefe, 2022), Mali (Armah-Ansah, 2023), and China (Wu et al., 2020). Pregnancy increases iron demands for both the mother and fetus, as the body requires more iron to produce additional blood to supply oxygen to the developing child. Furthermore, the overall blood volume expands during pregnancy, creating a higher demand for iron and vitamins to produce red blood cells, which naturally lowers hemoglobin (Hb) levels (Manickavasagam 2021, “Anaemia in Pregnancy: ACOG Practice Bulletin, Number 233,” 2021; James, 2021)

Women in rural areas had lower odds of severe anaemia compared to their urban counterparts, consistent with findings from East Africa (Teshale et al., 2020), and Sub-Saharan Africa (Moschovis et al. 2018, Worku et al., 2022). This reduced risk among rural women may stem from their better access to iron-rich foods, including dairy, eggs, and leafy greens, due to their agricultural activities. Rural women may also have greater access to fresh, locally grown produce, contributing to a more varied and nutritious diet that lowers anaemia risk (Tirore et al., 2023).

Women living in households with unimproved toilet facilities were found to have higher odds of severe anaemia compared to those in households with improved facilities. Similar findings were observed in studies from East Africa (Teshale et al., 2020), Lao (Keokenchanh et al., 2021), and Sub-Saharan Africa (Wu et al., 2020). This could be because women in unsanitary environments face a higher risk of food and waterborne illnesses, such as hookworm and diarrhea, which are known to contribute to anaemia. Families with unimproved toilets often have contaminated water sources, increasing the likelihood of gastrointestinal infections from bacteria, parasites, and other pathogens. Chronic infections like these can impair nutrient absorption, particularly of iron, which is vital in preventing anaemia (Yesuf et al., 2019).

Women in households using solid cooking fuels had a higher likelihood of severe anaemia compared to those using cleaner fuel sources. Studies from Ethiopia (Kanno et al., 2022), Sri Lanka (He et al., 2022), and China (Pathirathna et al., 2022), support these findings. This may be due to pollutants from solid fuel combustion leading to systemic inflammation, which impairs red blood cell production and disrupts iron balance, resulting in lower serum iron and anaemia. Additionally, indoor air pollution from burning solid fuels increases carbon monoxide exposure, which can inhibit iron absorption in the gastrointestinal tract, further contributing to iron-deficiency anaemia (Gillespie et al., 2010, Zou et al., 2014).

5.1. Strength and limitation of the study

Strength of the Study

This study used analytical type of study design which is cross sectional study design to assess spatial distribution of anaemia among women reproductive age in Ethiopia which needs less cost and time but relate cause and effect. This study also focus on determinant factors of anaemia among women reproductive age which get less attention by health care workers and the study population itself as the disease is chronic, but leads to devastating problem that can lead women reproductive age to death. Therefore identifying the risk factors by using such study help to plan prevention and control method from the grass route level before the problem occurred.

Limitation of the Study

This study has several limitations that may impact the interpretation of results. The design restricts the ability to determine cause-and-effect or temporal relationships between anaemia and

its associated risk factors. Additionally, using secondary data from the EDHS limits the inclusion of important variables such as dietary diversity, HIV status, malaria, hookworm infection, menorrhagia, and chronic diseases, making it challenging to fully capture the independent contributions of these factors to anaemia. Furthermore, the EDHS relied on participant recall for certain past events, introducing potential recall bias and affecting the reliability of reported information. The study was facility based of women reproductive age who did not came to health facility might be systematically different from those who came to health facility.

CHAPTER SIX

6. CONCLUSION AND RECOMMENDATION

6.1. CONCLUSION

This study aimed to assess the spatial distribution and determinants of anaemia among women of reproductive age in Ethiopia, employing various statistical methods, including global and local spatial statistics, as well as ordinal logistic regression (OLR) and generalized linear mixed models (GLMM). The findings reveal significant regional variation in anaemia prevalence, with Somali and Harari regions exhibiting the highest percentages of anaemia, while Addis Ababa, Tigray, SNNP, and Oromia recorded the lowest.

The spatial analysis demonstrated a clustered pattern of anaemia distribution, confirmed by the global Moran's I statistic (0.146652, p-value < 0.001). This clustering suggests that anaemia is not randomly distributed across the country but instead tends to group in certain regions. Further investigation using the Getis-Ord Gi* statistic revealed the presence of significant hot spots of anaemia in Somalia, Harari and Gambela regions, and cold spots in regions like Amhara, Oromia, SNNP, and Afar regions among WRA.

The ordinal logistic regression model identified several key predictors of anaemia severity, including marital status, number of living children, place of residence, region, wealth index, and religion. In particular, women from Amhara, SNNPR, and Addis Ababa had significantly lower odds of severe anaemia compared to those from Tigray. Age, urban residence, wealth, and religion also emerged as significant factors, indicating complex socio-demographic and environmental influences on anaemia status.

The results from the Brant test show significant violations of the proportional odds assumption for several key variables, indicating that their effects on the outcome are not consistent across categories. Variables such as region, age, type of cooking fuel, wealth index, number of living children, and marital status exhibit non-proportional effects, suggesting the need for alternative models like generalized ordinal logistic regression or multinomial logistic regression. On the other hand, variables like type of residence, source of drinking water, and pregnancy status meet the proportional odds assumption, with their effects remaining constant across outcome categories.

The application of the Generalized Linear Mixed Model (GLMM) provided a more nuanced understanding of the factors influencing anaemia status. Significant predictors include age, wealth index, religious affiliation, and household characteristics such as cooking fuel and refrigerator ownership. Importantly, the random effects for regions, while relatively small, indicate some unexplained variability in anaemia status across regions, likely reflecting differences in local conditions or unmeasured contextual factors.

Model diagnostics showed that the generalized linear mixed model (GLMM) offered a better fit compared to the ordinal logistic regression model, as indicated by the lower Bayesian Information Criterion (BIC) score. This model also accounted for the random effects associated with regional variability in anaemia prevalence, highlighting the importance of considering both fixed and random effects in understanding the distribution of anaemia.

In conclusion, the spatial and multivariable analyses underscore the critical role of regional, socio-economic, and demographic factors in shaping anaemia patterns among women of reproductive age in Ethiopia. Public health interventions should be tailored to address these disparities, particularly in regions identified as hot spots, and focus on improving nutrition, healthcare access, and socio-economic conditions to mitigate the burden of anaemia in the population.

6.2. RECOMMENDATIONS

Targeted Interventions for Higher-Risk Age Groups: Given that anaemia severity increases with age in women aged 20-29, focused interventions such as regular screenings and iron supplementation programs for this age group are recommended. Efforts to improve awareness of anaemia risks and preventative care are crucial, particularly for younger women entering this age range.

Enhanced Focus on Urban Areas: Since urban residence is associated with higher anaemia levels, it is essential to investigate and address urban-specific factors contributing to anaemia. Access to affordable healthcare, nutrition-focused education, and regular health screenings should be prioritized in urban areas, especially where healthcare disparities may exist.

Socio-Cultural Considerations in Anaemia Prevention: Religious affiliation has a varied impact on anaemia status. Public health campaigns could benefit from culturally sensitive outreach and engagement that takes these socio-cultural variations into account, thereby promoting anaemia prevention in alignment with community values and beliefs.

Economic Empowerment and Nutritional Support: Wealthier individuals experience higher anaemia levels, possibly due to lifestyle factors. Health education initiatives should target wealthier households, promoting balanced diets and lifestyle modifications that support iron-rich nutrition and overall health.

Cleaner Cooking Solutions: The positive impact of clean cooking fuels on reducing anaemia underscores the importance of promoting cleaner fuel alternatives. Public health campaigns should advocate for the adoption of cleaner cooking technologies, especially in areas where traditional solid fuels are still commonly used.

Regional Variability Considerations: Despite low regional variability, unexplained factors indicate the importance of region-specific interventions. Further research should aim to identify unique regional influences and tailor interventions to local contexts, addressing regional disparities in anaemia risk factors.

For Researcher: this was an institutional study further community based studies are recommended to identify risk factors for anaemia status in women reproductive age in Ethiopia.

For policy makers: needs a strengthen focus on pregnant women, women in households using solid cooking fuels and women who are in rural residence showed a severe anaemia status among women reproductive age in Ethiopia.

REFERENCES

Abioye, Abigail Adebisi, Christiana Adetoun Owopetu, Foyekemi Oyebola Adamu-Adedipe, Lydia Olubukola Odesanya, and Mary Ayodeji Olofin-Samuel. 2024. "Anaemia Prevention Among Pregnant Women." *International Journal of Nursing, Midwife and Health Related Cases* 10(2):1–11. doi: 10.37745/ijnmh.15/vol10n2111.

Agegenehu, Chilot Desta, Getayeneh Antehunegn Tesema, Achamyeleh Birhanu Teshale, Adugnaw Zeleke Alem, Yigizie Yeshaw, Sewnet Adem Kebede, and Alemneh Mekuriaw Liyew. 2021. "Spatial Distribution and Determinants of Iron Supplementation among Pregnant Women in Ethiopia: A Spatial and Multilevel Analysis." *Archives of Public Health* 79(1). doi: 10.1186/s13690-021-00669-2.

Aglago, Elom K., Edwige Landais, Francis Zotor, Genevieve Nicolas, Marc J. Gunter, Paul Amuna, and Nadia Slimani. 2018. "Optimising Design and Cost-Effective Implementation of Future Pan-African Dietary Studies: A Review of Existing Economic Integration and Nutritional Indicators for Scenario-Based Profiling and Clustering of Countries." in *Proceedings of the Nutrition Society*. Vol. 77.

Alem, Adugnaw Zeleke, Ferry Efendi, Lisa McKenna, Eva Belingon Felipe-Dimog, Dagmawi Chilot, Santo Imanuel Tonapa, Ika Adelia Susanti, and Agus Zainuri. 2023. "Prevalence and Factors Associated with Anemia in Women of Reproductive Age across Low- and Middle-Income Countries Based on National Data." *Scientific Reports* 13(1). doi: 10.1038/s41598-023-46739-z.

Andersen, Christopher T., Amare Worku Tadesse, Sabri Bromage, Habtamu Fekadu, Elena C. Hemler, Simone Passarelli, Donna Spiegelman, Christopher R. Sudfeld, Alemayehu Worku, Yemane Berhane, and Wafaie W. Fawzi. 2022. "Anemia Etiology in Ethiopia: Assessment of Nutritional, Infectious Disease, and Other Risk Factors in a Population-Based Cross-Sectional Survey of Women, Men, and Children." *Journal of Nutrition* 152(2). doi: 10.1093/jn/nxab366.

Andrade, Luís, and José Luís Silva. 2017. "Parameter Estimation for Weibull Distribution with Right Censored Data Using EM Algorithm." *Eksploatacja i Niezawodnosc* 19(2). doi: 10.17531/ein.2017.2.20.

Animut, Kassahun, and Getasew Berhanu. 2022. “Determinants of Anemia Status among Pregnant Women in Ethiopia: Using 2016 Ethiopian Demographic and Health Survey Data; Application of Ordinal Logistic Regression Models.” *BMC Pregnancy and Childbirth* 22(1). doi: 10.1186/s12884-022-04990-8.

Araújo-Pereira, Mariana, Virginia Sheikh, Irini Sereti, Beatriz Barreto-Duarte, María B. Arriaga, Rafael Tibúrcio, Caian L. Vinhaes, Manuella Pinto-de-Almeida, Jing Wang, Adam Rupert, Gregg Roby, Douglas Shaffer, Jintanat Ananworanich, Nittaya Phanuphak, Fred Sawe, and Bruno B. Andrade. 2022. “Association between Severe Anaemia and Inflammation, Risk of IRIS and Death in Persons with HIV: A Multinational Cohort Study.” *EBioMedicine* 85. doi: 10.1016/j.ebiom.2022.104309.

Armah-Ansah, Ebenezer Kwesi. 2023. “Determinants of Anemia among Women of Childbearing Age: Analysis of the 2018 Mali Demographic and Health Survey.” *Archives of Public Health* 81(1). doi: 10.1186/s13690-023-01023-4.

Ayele, Gizachew Sime, Abulie Takele Melku, and Semere Sileshi Belda. 2019. “Utilization of Skilled Birth Attendant at Birth and Associated Factors among Women Who Gave Birth in the Last 24 Months Preceding the Survey in Gura Dhamole Woreda, Bale Zone, Southeast Ethiopia.” *BMC Public Health* 19(1). doi: 10.1186/s12889-019-7818-6.

Chaparro, Camila M. an. Parminder S. Suchdev. 2019. “Annals NY Academy of Science - 2019 - Chaparro - Anemia Epidemiology Pathophysiology and Etiology in Low- and (1).Pdf.” *ANNALS OF THE NEW YORK ACADEMY OF SCIENCES* (Hemoglobin Concentration for Assessing Anemia).

Chaparro, Camila M., and Parminder S. Suchdev. 2019. “Anemia Epidemiology, Pathophysiology, and Etiology in Low- and Middle-Income Countries.” *Annals of the New York Academy of Sciences* 1450(1).

Cole, Stephen R., Haitao Chu, and Sander Greenland. 2014. “Maximum Likelihood, Profile Likelihood, and Penalized Likelihood: A Primer.” *American Journal of Epidemiology* 179(2). doi: 10.1093/aje/kwt245.

Czempik, Piotr F., and Agnieszka Wiórek. 2020. “Iron Deficiency in Sepsis Patients Managed

with Divided Doses of Iron Dextran: A Prospective Cohort Study." doi: 10.1038/s41598-023-32002-y.

Diggle, Peter. 2005. "Applied Spatial Statistics for Public Health Data." *Journal of the American Statistical Association* 100(470). doi: 10.1198/jasa.2005.s15.

Dolgun, Anil, and Osman Saracbasi. 2014. "Assessing Proportionality Assumption in the Adjacent Category Logistic Regression Model." *Statistics and Its Interface* 7(2). doi: 10.4310/SII.2014.v7.n2.a12.

Douthit, Nathan T., and Haimanot Kasahun Alemu. 2016. "Social Determinants of Health: Poverty, National Infrastructure and Investment." *BMJ Case Reports* 2016. doi: 10.1136/bcr-2016-215670.

Ejigu, Bedilu Alamirie, Eshetu Wencheko, and Kiros Berhane. 2018. "Spatial Pattern and Determinants of Anaemia in Ethiopia." *PLoS ONE* 13(5). doi: 10.1371/journal.pone.0197171.

Firth, David. 1987. "On the Efficiency of Quasi-Likelihood Estimation." *Biometrika* 74(2). doi: 10.1093/biomet/74.2.233.

Gautam, Sujan, Haju Min, Heenyun Kim, and Hyoung Sun Jeong. 2019. "Determining Factors for the Prevalence of Anemia in Women of Reproductive Age in Nepal: Evidence from Recent National Survey Data." *PLoS ONE* 14(6). doi: 10.1371/journal.pone.0218288.

Gebreegziabher, Tafere, and Saran Sidibe. 2023. "Prevalence and Contributing Factors of Anaemia among Children Aged 6-24 Months and 25-59 Months in Mali." *Journal of Nutritional Science* 12. doi: 10.1017/jns.2023.93.

Gebremedhin, Samson, and Fikre Enquselassie. 2011. "Correlates of Anemia among Women of Reproductive Age in Ethiopia: Evidence from Ethiopian DHS 2005." *Ethiopian Journal of Health Development* 25(1). doi: 10.4314/ejhd.v25i1.69842.

Gebrerufael, Gebru Gebremeskel, and Bsrat Tesfay Hagos. 2023. "Anemia Prevalence and Risk Factors in Two of Ethiopia's Most Anemic Regions among Women: A Cross-Sectional Study." *Advances in Hematology* 2023. doi: 10.1155/2023/2900483.

Geisser, Seymour. 1974. “A Predictive Approach to the Random Effect Model.” *Biometrika* 61(1). doi: 10.1093/biomet/61.1.101.

Gemechu, Kabtamu, Haftu Asmerom, Lealem Gedefaw, Mesay Arkew, Tilahun Bete, and Wondimagegn Adissu. 2023. “Anemia Prevalence and Associated Factors among School-Children of Kersa Woreda in Eastern Ethiopia: A Cross-Sectional Study.” *PLoS ONE* 18(3 March). doi: 10.1371/journal.pone.0283421.

Georgieff, Michael K., Nancy F. Krebs, and Sarah E. Cusick. 2019. “The Benefits and Risks of Iron Supplementation in Pregnancy and Childhood.” *Annual Review of Nutrition* 39.

Geta, Teshome Gensa, Samson Gebremedhin, and Akinyinka O. Omigbodun. 2022. “Prevalence and Predictors of Anemia among Pregnant Women in Ethiopia: Systematic Review and Meta-Analysis.” *PLoS ONE* 17(7 July).

Grueber, C. E., S. Nakagawa, R. J. Laws, and I. G. Jamieson. 2011. “Multimodel Inference in Ecology and Evolution: Challenges and Solutions.” *Journal of Evolutionary Biology* 24(4).

Guan, Yongtao, Abdollah Jalilian, and Rasmus Waagepetersen. 2015. “Quasi-Likelihood for Spatial Point Processes.” *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 77(3). doi: 10.1111/rssb.12083.

Hasan, Md Mehedi, Ricardo J. Soares Magalhaes, Sarah P. Garnett, Yaqoot Fatima, Md Tariqujjaman, Sonia Pervin, Saifuddin Ahmed, and Abdullah A. Mamun. 2022. “Anaemia in Women of Reproductive Age in Low-and Middle-Income Countries: Progress towards the 2025 Global Nutrition Target.” *Bulletin of the World Health Organization* 100(3). doi: 10.2471/BLT.20.280180.

He, Yaling, Xiaotian Liu, Yiquan Zheng, Zhihan Zhai, Xueyan Wu, Ning Kang, Yinghao Yuchi, Jian Hou, Zhenxing Mao, Wenqian Huo, Chongjian Wang, and Yuqian Li. 2022. “Lower Socioeconomic Status Strengthens the Effect of Cooking Fuel Use on Anemia Risk and Anemia-Related Parameters: Findings from the Henan Rural Cohort.” *Science of the Total Environment* 831. doi: 10.1016/j.scitotenv.2022.154958.

Hoenemann, Christian, Norbert Ostendorf, Alexander Zarbock, Dietrich Doll, Olaf Hagemann,

Mathias Zimmermann, and Markus Luedi. 2021. “Reticulocyte and Erythrocyte Hemoglobin Parameters for Iron Deficiency and Anemia Diagnostics in Patient Blood Management. A Narrative Review.” *Journal of Clinical Medicine* 10(18).

Hounton, Sennen, William Winfrey, Aluisio J. D. Barros, and Ian Askew. 2015. “Patterns and Trends of Postpartum Family Planning in Ethiopia, Malawi, and Nigeria: Evidence of Missed Opportunities for Integration.” *Global Health Action* 8(1). doi: 10.3402/gha.v8.29738.

Kanno, Girum Gebremeskel, Temesgen Geremew, Tesfaye Diro, Stephen Vincent Musarapasi, Renay Van Wyk, Binyam Tariku Seboka, Awash Alembo, Robel Hussen, Negasa Eshete Soboksa, and Mekonnen Birhane Aregu. 2022. “The Link between Indoor Air Pollution from Cooking Fuels and Anemia Status among Non-Pregnant Women of Reproductive Age in Ethiopia.” *SAGE Open Medicine* 10. doi: 10.1177/20503121221107466.

Keokenchanh, Sengtavanh, Sengchanh Kounnavong, Akiko Tokinobu, Kaoru Midorikawa, Wakaha Ikeda, Akemi Morita, Takumi Kitajima, and Shigeru Sokejima. 2021. “Prevalence of Anemia and Its Associate Factors among Women of Reproductive Age in Lao PDR: Evidence from a Nationally Representative Survey.” *Anemia* 2021. doi: 10.1155/2021/8823030.

Kibret, Kelemu Tilahun, Catherine Chojenta, Ellie D’Arcy, and Deborah Loxton. 2019. “Spatial Distribution and Determinant Factors of Anaemia among Women of Reproductive Age in Ethiopia: A Multilevel and Spatial Analysis.” *BMJ Open* 9(4):1–14. doi: 10.1136/bmjopen-2018-027276.

Lelisho, Mesfin Esayas, Abebe Argaw Wogi, and Seid Ali Tareke. 2022. “Ordinal Logistic Regression Analysis in Determining Factors Associated with Socioeconomic Status of Household in Tepi Town, Southwest Ethiopia.” *Scientific World Journal* 2022. doi: 10.1155/2022/2415692.

Lingani, Moussa, Serge Henri Zango, Innocent Valéa, Sékou Samadoulougou, Moussa Abdel Sanou, Hermann Sorgho, Edmond Sawadogo, Michèle Dramaix, Philippe Donnen, Annie Robert, and Halidou Tinto. 2024. “Prevalence and Determinants of Anaemia among

Pregnant Women in a High Malaria Transmission Setting: A Cross-Sectional Study in Rural Burkina Faso." *Pan African Medical Journal* 47. doi: 10.11604/pamj.2024.47.2.40612.

Liu, Anqi, Hua He, Xin M. Tu, and Wan Tang. 2023. "On Testing Proportional Odds Assumptions for Proportional Odds Models." *General Psychiatry* 36(3).

Luque-Fernandez, Miguel Angel, Michael Schomaker, Bernard Rachet, and Mireille E. Schnitzer. 2018. "Targeted Maximum Likelihood Estimation for a Binary Treatment: A Tutorial." *Statistics in Medicine* 37(16). doi: 10.1002/sim.7628.

Manickavasagam, Madhubala. 2021. "Haemoglobin Level of Pregnant Women on First Appointment to Antenatal Care Clinic and Their Awareness on Anemia during Pregnancy." *Journal of Obstetrics, Gynecology and Cancer Research* 6(2). doi: 10.30699/jogcr.6.2.57.

Mawani, Minaz, and Savera Aziz Ali. 2016. "Iron Deficiency Anemia among Women of Reproductive Age, an Important Public Health Problem: Situation Analysis." *Reproductive System & Sexual Disorders* 5(3). doi: 10.4172/2161-038x.1000187.

McCulloch, Charles E. 1996. "AN INTRODUCTION TO GENERALIZED LINEAR MIXED MODELS." *Conference on Applied Statistics in Agriculture*. doi: 10.4148/2475-7772.1314.

McLean, Erin, Mary Cogswell, Ines Egli, Daniel Wojdyla, and Bruno De Benoist. 2009. "Worldwide Prevalence of Anaemia, WHO Vitamin and Mineral Nutrition Information System, 1993-2005." *Public Health Nutrition* 12(4). doi: 10.1017/S1368980008002401.

McNeish, Daniel. 2020. "Should We Use F-Tests for Model Fit Instead of Chi-Square in Overidentified Structural Equation Models?" *Organizational Research Methods* 23(3). doi: 10.1177/1094428118809495.

Mehrolia, Sangeeta, Subburaj Alagarsamy, and Vijay Mallikraj Solaikutty. 2021. "Customers Response to Online Food Delivery Services during COVID-19 Outbreak Using Binary Logistic Regression." *International Journal of Consumer Studies* 45(3). doi: 10.1111/ijcs.12630.

Moschovis, Peter P., Matthew O. Wiens, Lauren Arlington, Olga Antsygina, Douglas Hayden, Walter Dzik, Julius P. Kiwanuka, David C. Christiani, and Patricia L. Hibberd. 2018.

"Individual, Maternal and Household Risk Factors for Anaemia among Young Children in Sub-Saharan Africa: A Cross-Sectional Study." *BMJ Open* 8(5). doi: 10.1136/bmjopen-2017-019654.

Murdock, John E., Nicholas D. K. Petraco, John I. Thornton, Michael T. Neel, Todd J. Weller, Robert M. Thompson, James E. Hamby, and Eric R. Collins. 2017. "The Development and Application of Random Match Probabilities to Firearm and Toolmark Identification." *Journal of Forensic Sciences* 62(3). doi: 10.1111/1556-4029.13386.

Mwangi, Martin N., Glory Mzembe, Ernest Moya, and Hans Verhoef. 2021. "Iron Deficiency Anaemia in Sub-Saharan Africa: A Review of Current Evidence and Primary Care Recommendations for High-Risk Groups." *The Lancet Haematology* 8(10).

Negash, Berhan Tsegaye, and Mohammed Ayalew. 2023. "Trend and Factors Associated with Anemia among Women Reproductive Age in Ethiopia: A Multivariate Decomposition Analysis of Ethiopian Demographic and Health Survey." *PLoS ONE* 18(1 January). doi: 10.1371/journal.pone.0280679.

Nti, Josephine, Seth Afagbedzi, Frances Baaba Da-Costa Vroom, Noor Akma Ibrahim, and Chris Guure. 2021. "Variations and Determinants of Anemia among Reproductive Age Women in Five Sub-Saharan Africa Countries." *BioMed Research International* 2021. doi: 10.1155/2021/9957160.

Oliveira, Fernando L. P., André L. F. Cançado, Gustavo de Souza, Gladston J. P. Moreira, and Martin Kulldorff. 2018. "Border Analysis for Spatial Clusters." *International Journal of Health Geographics* 17(1). doi: 10.1186/s12942-018-0124-1.

Owais, Aatekah, Catherine Merritt, Christopher Lee, and Zulfiqar A. Bhutta. 2021. "Anemia among Women of Reproductive Age: An Overview of Global Burden, Trends, Determinants, and Drivers of Progress in Low-and Middle-Income Countries." *Nutrients* 13(8). doi: 10.3390/nu13082745.

Pan, Zhiying, and D. Y. Lin. 2005. "Goodness-of-Fit Methods for Generalized Linear Mixed Models." *Biometrics* 61(4). doi: 10.1111/j.1541-0420.2005.00365.x.

Pathirathna, Malshani Lakshika, Buddhini Piumi Pabasara Samarasekara, Charitha Mendis, Chandraratne Mahinda Bandara Dematawewa, Kayako Sekijima, Mieko Sadakata, Yoshiyuki Muramatsu, and Naoshi Fujiwara. 2022. "Is Biomass Fuel Smoke Exposure Associated with Anemia in Non-Pregnant Reproductive-Aged Women?" *PLoS ONE* 17(8 August). doi: 10.1371/journal.pone.0272641.

Payne, Elizabeth H., Mulugeta Gebregziabher, James W. Hardin, Viswanathan Ramakrishnan, and Leonard E. Egede. 2018. "An Empirical Approach to Determine a Threshold for Assessing Overdispersion in Poisson and Negative Binomial Models for Count Data." *Communications in Statistics: Simulation and Computation* 47(6). doi: 10.1080/03610918.2017.1323223.

Prasetyo, Rindang Bangun, Heri Kuswanto, Nur Iriawan, and Brodjol Sutijo Suprih Ulama. 2020. "Binomial Regression Models with a Flexible Generalized Logit Link Function." *Symmetry* 12(2). doi: 10.3390/sym12020221.

Putra, Rizqi Widyantori Hasanah, J. Supadi, and Wiwik Wijaningsih. 2019. "PENGARUH PEMBERIAN EDUKASI GIZI TERHADAP PENGETAHUAN DAN SIKAP MENGENAI ANEMIA PADA REMAJA PUTRI." *JURNAL RISET GIZI* 7(2). doi: 10.31983/jrg.v7i2.5220.

Raiten, Daniel J., Sorrel Namasté, and Bernard Brabin. 2011. "Considerations for the Safe and Effective Use of Iron Interventions in Areas of Malaria Burden - Executive Summary." *International Journal for Vitamin and Nutrition Research* 81(1). doi: 10.1024/0300-9831/a000051.

Saad, Shihab Elbagir, and Jing Yang. 2019. "Twitter Sentiment Analysis Based on Ordinal Regression." *IEEE Access* 7. doi: 10.1109/ACCESS.2019.2952127.

Safiri, Saeid, Ali Asghar Kolahi, Maryam Noori, Seyed Aria Nejadghaderi, Nahid Karamzad, Nicola Luigi Bragazzi, Mark J. M. Sullman, Morteza Abdollahi, Gary S. Collins, Jay S. Kaufman, and Jessica A. Grieger. 2021. "Burden of Anemia and Its Underlying Causes in 204 Countries and Territories, 1990–2019: Results from the Global Burden of Disease Study 2019." *Journal of Hematology and Oncology* 14(1). doi: 10.1186/s13045-021-01202-

2.

Sahile, Abay, Dereje Bekele, and Habtamu Ayele. 2022. “Determining Factors of Neonatal Mortality in Ethiopia: An Investigation from the 2019 Ethiopia Mini Demographic and Health Survey.” *PLoS ONE* 17(12 December). doi: 10.1371/journal.pone.0267999.

Saini, Kamal, Rajveer Singh, and Divyani Agrawal. 2019. “Comparing the Safety between Ferrous Carboxymaltose and Iron Sucrose Therapy in Iron Deficiency Anemia during Pregnancy in Obstetrics and Gynecology Ward at Tertiary Care Hospital, Jaipur.” *Indian Journal of Pharmacy Practice* 12(3). doi: 10.5530/ijopp.12.3.37.

Sappani, Marimuthu, Thenmozhi Mani, Edwin Sam Asirvatham, Melvin Joy, Malavika Babu, and Lakshmanan Jeyaseelan. 2023. “Trends in Prevalence and Determinants of Severe and Moderate Anaemia among Women of Reproductive Age during the Last 15 Years in India.” *PLoS ONE* 18(6 JUNE). doi: 10.1371/journal.pone.0286464.

Sawada M. 2001. “Global Spatial Autocorrelation Indices- Moran’s I, Geary’s C and the General Cross-Product Statistic.” *Department of Geography, University of Ottawa*.

Seltman, Howard J. 2014. “Mixed Models: A Flexible Approach to Correlated Data.” in *Experimental Design and Analysis*.

Semedo, Rosa M. L., Marta M. A. S. Santos, Mirian R. Baião, Ronir R. Luiz, and Gloria V. Da Veiga. 2014. “Prevalence of Anaemia and Associated Factors among Children below Five Years of Age in Cape Verde, West Africa.” *Journal of Health, Population and Nutrition* 32(4).

Sharma, Andrea J., O. Yaw Addo, Zuguo Mei, and Parminder S. Suchdev. 2019. “Reexamination of Hemoglobin Adjustments to Define Anemia: Altitude and Smoking.” *Annals of the New York Academy of Sciences* 1450(1). doi: 10.1111/nyas.14167.

Sharma, Jyoti, Sriram Devanathan, Angan Sengupta, and P. N. Rajeshwari. 2024. “Assessing the Prevalence of Iron Deficiency Anemia and Risk Factors among Children and Women: A Case Study of Rural Uttar Pradesh.” *Clinical Epidemiology and Global Health* 26. doi: 10.1016/j.cegh.2024.101545.

Shitu, Kegnie, and Bewuketu Terefe. 2022. “Anaemia and Its Determinants among Reproductive Age Women (15–49 Years) in the Gambia: A Multi-Level Analysis of 2019–20 Gambian Demographic and Health Survey Data.” *Archives of Public Health* 80(1). doi: 10.1186/s13690-022-00985-1.

Signorovitch, J., R. Ayyagari, D. Cheng, and E. Q. Wu. 2013. “Matching-Adjusted Indirect Comparisons: A Simulation Study of Statistical Performance.” *Value in Health* 16(3). doi: 10.1016/j.jval.2013.03.271.

Takele, Tesfaye B., and Tassew D. Tolcha. 2021. “Optimal Transit Corridors for Ethiopia.” *Journal of Transport and Supply Chain Management* 15. doi: 10.4102/jtscm.v15i0.567.

Tesema, Getayeneh Antehunegn, Misganaw Gebrie Worku, Zemenu Tadesse Tessema, Achamyeleh Birhanu Teshale, Adugnaw Zeleke Alem, Yigizie Yeshaw, Tesfa Sewunet Alamneh, and Alemneh Mekuriaw Liyew. 2021. “Prevalence and Determinants of Severity Levels of Anemia among Children Aged 6-59 Months in Sub-Saharan Africa: A Multilevel Ordinal Logistic Regression Analysis.” *PLoS ONE* 16(4 April). doi: 10.1371/journal.pone.0249978.

Teshale, Achamyeleh Birhanu, Getayeneh Antehunegn Tesema, Misganaw Gebrie Worku, Yigizie Yeshaw, and Zemenu Tadesse Tessema. 2020. “Anemia and Its Associated Factors among Women of Reproductive Age in Eastern Africa: A Multilevel Mixed-Effects Generalized Linear Model.” *PLoS ONE* 15(9 September). doi: 10.1371/journal.pone.0238957.

Tirore, Lire Lemma, Abriham Shiferaw Areba, Aklilu Habte, Mitiku Desalegn, and Abraham Sahilemichael Kebede. 2023. “Prevalence and Associated Factors of Severity Levels of Anemia among Women of Reproductive Age in Sub-Saharan Africa: A Multilevel Ordinal Logistic Regression Analysis.” *Frontiers in Public Health* 11. doi: 10.3389/fpubh.2023.1349174.

Tsegaye, Berhan, Belay Amare, and Mulu Reda. 2021. “Prevalence and Factors Associated with Immediate Postnatal Care Utilization in Ethiopia: Analysis of Ethiopian Demographic Health Survey 2016.” *International Journal of Women’s Health* 13. doi:

10.2147/IJWH.S294058.

Tusa, Biruk Shalmeno, Sewnet Adem Kebede, and Adisu Birhanu Weldesenbet. 2021. “Spatial Distribution and Determinant Factors of Anemia among Adults Aged 15–59 in Ethiopia; Using Mixed-Effects Ordinal Logistic Regression Model.” *BMC Nutrition* 7(1). doi: 10.1186/s40795-021-00424-4.

Umeta, Melaku, Jemal Haidar, Tsegaye Demissie, Girma Akalu, and Gonfa Ayana. 2008. “Iron Deficiency Anaemia among Women of Reproductive Age in Nine Administrative Regions of Ethiopia.” *Ethiopian Journal of Health Development*.

Wasserman, Larry. 2000. “Bayesian Model Selection and Model Averaging.” *Journal of Mathematical Psychology* 44(1). doi: 10.1006/jmps.1999.1278.

Wedderburn, R. W. M. 1974. “Quasi Likelihood Functions, Generalized Linear Models, and the Gauss Newton Method.” *Biometrika* 61(3). doi: 10.2307/2334725.

WHO. 2017. “Reproductive Health Indicators: Guidelines for Their Generation, Interpretation and Analysis for Global Monitoring.” *Geneva: World Health Organization* 57(2).

WHO. 2021. “Prevalence of Anemia and Its Associated Factors among Under-Five Age Children in Shanan Gibe Hospital, Southwest Ethiopia.” *BMC Pediatrics* 21(1). doi: 10.1186/s12887-021-03011-5.

Williams, Anne M., Kenneth H. Brown, Lindsay H. Allen, Omar Dary, Denish Moorthy, and Parminder S. Suchdev. 2023. “Improving Anemia Assessment in Clinical and Public Health Settings.” *Journal of Nutrition* 153. doi: 10.1016/j.tjnut.2023.05.032.

Woldu, Berhanu, Bamlaku Enawgaw, Fikir Asrie, Elias Shiferaw, Zegeye Getaneh, and Mulugeta Melku. 2020. “Prevalence and Associated Factors of Anemia among Reproductive-Aged Women in Sayint Adjibar Town, Northeast Ethiopia: Community-Based Cross-Sectional Study.” *Anemia* 2020. doi: 10.1155/2020/8683946.

World Health Organization Team. 2014. *Global Nutrition Targets 2025 Stunting Policy Brief*.

Wu, Yu, Hanfeng Ye, Jihong Liu, Qiuyue Ma, Yanling Yuan, Qian Pang, Jue Liu, Cai Kong, and

Min Liu. 2020. “Prevalence of Anemia and Sociodemographic Characteristics among Pregnant and Non-Pregnant Women in Southwest China: A Longitudinal Observational Study.” *BMC Pregnancy and Childbirth* 20(1). doi: 10.1186/s12884-020-03222-1.

Xu, Haiqi, Enkhbold Nyamsuren, Simon Scheider, and Eric Top. 2023. “A Grammar for Interpreting Geo-Analytical Questions as Concept Transformations.” *International Journal of Geographical Information Science* 37(2). doi: 10.1080/13658816.2022.2077947.

Yadav, Jeetendra, and Nilima Nilima. 2021. “Geographic Variation and Factors Associated with Anemia among Under-Fives in India: A Multilevel Approach.” *Clinical Epidemiology and Global Health* 9. doi: 10.1016/j.cegh.2020.09.008.

Yesuf, Dejene Abraham, Lensa Tesfaye Abdissa, Emiru Adeba Gerbi, and Edosa Kifle Tola. 2019. “Prevalence of Intestinal Parasitic Infection and Associated Factors among Pregnant Women Attending Antenatal Care at Public Health Facilities in Lalo Kile District, Oromia, Western Ethiopia.” *BMC Research Notes* 12(1). doi: 10.1186/s13104-019-4781-3.

Zanakis, Stelios H., and Jerzy Kyparisis. 1986. “A Review of Maximum Likelihood Estimation Methods for the Three-Parameter Weibull Distribution.” *Journal of Statistical Computation and Simulation* 25(1–2). doi: 10.1080/00949658608810924.

Zelelew, Hailu. 2012. *Health Care Financing Reform in Ethiopia: Improving Quality and Equity*.

Zou, Ying Ying, Yun Yuan, Enci M. Kan, Jia Lu, and Eng Ang Ling. 2014. “Combustion Smoke-Induced Inflammation in the Olfactory Bulb of Adult Rats.” *Journal of Neuroinflammation* 11(1). doi: 10.1186/s12974-014-0176-5.

APPENDICES

Table 3.1: Variables coding

SN.	Variable	Code
1	Anaemia Status	1 = no, 2 = mild, 3 = moderate and 4 = severe
2	Region	Tigray = 1, Afar = 2, Amhara = 3, Oromia = 4, Somali = 5, Benishangul = 6, SNNPR = 7, Gambela = 8, Harari = 9, Addis Ababa = 10 and Dire Dawa = 11
3	Age in 5 year groups	15-19 =1, 20-24 =2, 25-29 =3, 30-34 =4, 35-39 =5, 40-44 =6 and 45-49 =7
4	Type of place of residence	Rural = 1 and Urban = 2
5	Religion	Orthodox = 1, Catholic = 2, protestant = 3, Muslim = 4, Traditional = 5 and Other
6	Type of toilet facility	No facility = 1, Toilet with flush = 2 and Fit latrine= 3
7	Educational attainment	No education = 0, Incomplete primary = 1, Complete primary = 2, Incomplete secondary = 3, Complete secondary = 4 and Higher = 5
8	Wealth index combined	Poorest = 1, Poorer = 2, Middle = 3, Richer = 4 and Richest = 5
9	Source of drinking water	Protected = 1 and Unprotected = 2
10	Type of cooking fuel	Clean =1 and Solid = 2
11	Pregnancy status	No = 1 and Yes = 2
13	Use of contraception	No = 1 and Yes = 2
14	Marital status	Never in union = 1, Married = 2, Divirced = 3 and

		Widowed
15	Availability of electricity	No = 1 and Yes = 2
16	Household has radio	No = 1 and Yes = 2
17	Household has Tv	No = 1 and Yes = 2
18	Household has refrigerator	No = 1 and Yes = 2
19	Latnum	Defined in Number
20	Longnum	Defined in Number
21	Number of living children	Defined in Number

Table 4.3: Prevalence of anaemia status by predictors in reproductive age women

Predictors	Anaemia status in women reproductive age			
	No anaemia	Mild	Moderate	Severe
Age in 5 year groups				
15-19	59.2	18.0	3.62	19.1
20-24	57.0	15.8	10.1	17.0
25-29	56.6	18.3	8.39	16.7
30-34	54.5	21.6	6.09	17.8
35-39	57.2	24.3	3.66	14.9
40-44	58.7	24.8	2.94	13.6
45-49	57.1	27.7	3.90	11.3
Type of place of residence				
Urban	56.1	14.7	10.7	18.4
Rural	57.9	22.6	3.66	15.8
Religion				
Orthodox	65.1	21.3	6.58	6.99
Catholic	47.4	24.4	10.3	17.9

Protestant	62.0	24.5	7.07	6.37
Muslim	48.0	16.2	4.95	30.8
Traditional	61.7	38.3	-	-
Other	44.4	25.9	11.1	18.5
Educational attainment				
No education	54.5	23.6	1.40	20.5
Incomplete primary	55.4	32.0	0.881	11.8
Complete primary	99.4	-	0.197	0.394
Incomplete secondary	65.3	-	0.0981	34.6
Complete secondary	58.5	1.54	37.7	2.31
Higher	38.5	1.46	54.2	5.86
Wealth index combined				
Poorest	50.2	20.6	2.51	26.7
Poorer	60.3	26.0	1.72	12.1
Middle	62.5	25.1	3.39	8.99
Richer	59.5	23.5	4.76	12.2
Richest	57.5	13.0	12.2	17.3
Source of drinking water				
Protected	56.9	20.9	7.17	15.0
Un protected	58.2	17.8	3.12	20.9
Pregnancy status				
No	57.7	20.1	6.09	16.0
Yes	52.8	18.6	5.23	23.3
Use of contraception				
No	56.0	19.2	5.81	19.0
Yes	61.5	22.6	6.66	9.26
Marital status				
Never in union	59.8	15	7.70	17.5

Married	57.0	21.0	5.38	16.6
Divorced	54.2	24.2	6.82	14.8
Widowed	48.5	33.9	2.64	15.0
Household has radio				
No	56.4	20.7	5.32	17.6
Yes	59.4	18.3	7.63	14.7
Household has Tv				
No	57.2	22.9	3.76	16.1
Yes	57.6	12.0	12.2	18.2
Household has refrigerator				
No	56.8	21.8	4.38	17.1
Yes	60	10.4	14.8	14.7

Table 4.5: Bi-variable and multivariable mixed-effect ordinal logistic regression model of anaemia among reproductive age of women in Ethiopia

Variables	Crude odds ratio			Adjusted odds ratio		
	OR.	95% CI		OR.	95% CI	p-value
Region						
Tigray	Ref	Ref .		Ref	Ref	Ref
Afar	.425	[0.333, 0.542]		1.057	[0..79, 1.416]	.709
Amhara	.462	[0.369, 0.578]		**.683**	**[0.539,.866]**	**.002**
Oromia	.476	[0.382, 0.592]		.791	[0.615, 1.016]	.067
Somali	.385	[0.301, 0.493]		1.026	[0.764, 1.38]	.863
Benishangul	.577	[0.457, 0.727]		1.022	[0.79, 1.323]	.866
SNNPR	.511	[0.41, 0.635]		**.743**	**[0.568, 0.971]**	**.03**
Gambela	.802	[0.636, 1.012]		1.081	[0.826, 1.415]	.57
Harari	1.182	[0.95, 1.471]		.992	[0.771, 1.277]	.952
AddisAbaba	2.272	[1.849, 2.79]		**.707**	**[0.553, 0.904]**	**.006**
Dire Dawa	1.223	[0.985, 1.517]		.976	[0.759, 1.256]	.852
Age in 5-yr groups						
15-19	Ref	Ref		Ref	Ref	Ref
20-24	1.509	[1.306, 1.743]		**1.648**	**[1.401, 1.939]**	**0**
25-29	1.169	[1.015, 1.347]		**1.678**	**[1.403, 2.007]**	**0**
30-34	.818	[0.695, 0.962]		**1.392**	**[1.127, 1.719]**	**.002**
35-39	.684	[0.577, 0.811]		1.256	[0.999, 1.578]	.051
40-44	.645	[0.531, 0.783]		1.272	[0.981, 1.648]	.069

45-49	.636	[0.512, 0.789]		1.305	[0.984, 1.731]	.065
Type of place of residence						
Rural	Ref	Ref		Ref	Ref	Ref
Urban	3.734	[3.377, 4.129]		**1.22**	**[1.025, 1.452]**	**.025**
Religion						
Orthodox	Ref	Ref		Ref	Ref	Ref
Catholic	1.462	[0.933, 2.291]		**1.997**	**[1.232, 3.235]**	**.005**
Protestant	.74	[0.651, 0.841]		1.105	[0.936, 1.305]	.237
Muslim	.454	[0.408, 0.505]		**.604**	**[0.522, 0.699]**	**0**
Traditional	.292	[0.163, 0.522]		.578	[0.306, 1.09]	.09
Other	1.803	[0.876, 3.711]		**3.069**	**[1.419, 6.637]**	**.004**
Type of toilet facility						
No facility	Ref	Ref .		Ref	Ref	Ref
Toilet with flush	6.01	[4.986, 7.245]		1.148	[0.898, 1.468]	.271
Fit latrine	1.798	[1.618, 1.997]		.971	[0.843, 1.119]	.686
Wealth index combined						
Poorest	Ref	Ref .		Ref	Ref	Ref
Poorer	1.093	[0.923, 1.294]		1.029	[0.851, 1.244]	.768
Middle	1.365	[1.149, 1.622]		**1.274**	**[1.041, 1.56]**	**.019**
Richer	1.695	[1.43, 2.009]		**1.579**	**[1.27, 1.963]**	**0**
Richest	5.407	[4.718, 6.196]		**2.409**	**[1.798, 3.228]**	**0**
Source of drinking water						
Protected	Ref	Ref .		Ref	Ref	Ref
Unprotected	.493	[0.442, 0.549]		.923	[0.813, 1.048]	.217
Type of cooking fuel						
Clean	Ref	Ref .		Ref	Ref	Ref
Solid	.198	[0.177, 0.222]		**.541**	**[0.458, 0.639]**	**0**
Pregnancy status						
No	Ref	Ref .		Ref	Ref	Ref
Yes	.846	[0.718, 0.998]		1.026	[0.852, 1.236]	.787
Number of living child	.823	[0.807, 0.84]		**.923**	**[0.894, 0.954]**	**0**
Use of contraception						
No	Ref	Ref .		Ref	Ref	Ref
Yes	1.039	[0.933, 1.158]		1.022	[.896, 1.165]	.747
Marital status						
Never in union	Ref	Ref .		Ref	Ref	Ref
Married	.445	[0.4, .495]		**.636**	**[0.541, 0.747]**	**0**
Divorced	.512	[0.42, .625]		**.501**	**[0.401, 0.625]**	**0**
Widowed	.381	[0.281, .518]		**.548**	**[0.388, 0.773]**	**.001**
Availability of electricity						

No	Ref	Ref		Ref	Ref	Ref
Yes	3.232	[2.928, 3.568]		**.799**	**[0.662, 0.964]**	**.019**
Household has radio						
No	Ref	Ref		Ref	Ref	Ref
Yes	1.604	[1.452, 1.771]		1.032	[0.923, 1.154]	.579
Household has Tv						
No	Ref	Ref .		Ref	Ref	Ref
Yes	4.039	[3.643, 4.478]		1.133	[0.921, 1.394]	.237
Household has refrigerator						
No	Ref	Ref		Ref	Ref	Ref
Yes	4.497	[4.01, 5.043]		**1.317**	**[1.105, 1.568]**	**.002**

Table 4.6: Brant test of parallel regression assumption

Variables	chi^2	$p>chi^2$
All	2354.430	0.000
Region		
Afar	34.460	0.000
Amhara	6.590	0.037
Oromia	8.010	0.018
Somali	7.640	0.022
Benishangul	19.630	0.000
SNNPR	9.340	0.009
Gambela	13.360	0.001
Harari	32.000	0.000
AddisAbaba	29.720	0.000
Dire Dawa	23.590	0.000
Age in five year groups		
20-24	184.900	0.000
25-29	161.940	0.000
30-34	110.800	0.000
35-39	75.870	0.000
40-44	54.080	0.000

45-49	46.650	0.000
Type of place of residence		
Urban	2.860	0.239
Type of toilet facility		
Toilet with flush	4.730	0.094
Fit latrine	6.280	0.043
Wealth index combined		
Poorer	14.330	0.001
Middle	56.210	0.000
Richer	101.110	0.000
Richest	109.540	0.000
Source of drinking water		
Unprotected	2.470	0.291
Type of cooking fuel		
Solid	47.460	0.000
Pregnancy status		
Yes	0.830	0.659
Number of living children	206.740	0.000
Use of contraception		
Yes	6.100	0.047
Marital status		
Married	7.760	0.021
Divorced	18.230	0.000
Widowed	5.770	0.056
Availability of electricity		
Yes	0.450	0.799
House hold has radio		
Yes	4.270	0.118
House hold has Tv		

Yes	10.740	0.005
House hold has refrigerator		
Yes	33.350	0.000

Table 4.7: Generalized linear mixed-effect model of anemia status among reproductive age of women in Ethiopia

Anaemia status	Coefficient	Std. err	z	P>|z|	[95% conf. interval]	
Age in 5year groups						
2	.499938	**.0828349**	**6.04**	**0.000**	**.3375846**	**.6622914**
3	.5110133	**.0913009**	**5.60**	**0.000**	**.3320668**	**.6899597**
4	.3230739	**.1076253**	**3.00**	**0.003**	**.1121323**	**.5340156**
5	.2180075	**.116511**	**1.87**	**0.061**	**-.0103499**	**.4463649**
6	.2293134	**.1322033**	**1.73**	**0.083**	**-.0298004**	**.4884271**
7	.2506332	**.1439454**	**1.74**	**0.082**	**-.0314946**	**.5327611**
Type of place of residence						
	.205031	**.088133**	**2.33**	**0.020**	**.0322934**	**.3777685**
Religion						
2	.7107044	**.2442686**	**2.91**	**0.004**	**.23194671**	**1.189462**
3	.1044179	**.0794763**	**1.31**	**0.189**	**-.0513528**	**.2601885**
4	-.4634591	**.0701223**	**-6.61**	**0.000**	**-.6008963**	**-.326022**
5	-.5515398	**.3224042**	**-1.71**	**0.087**	**-1.18344**	**.0803607**
6	1.1307	**.3924543**	**2.88**	**0.004**	**.3615035**	**1.899896**
Type of toile facility						
2	.1046022	**.1247238**	**0.84**	**0.402**	**-.139852**	**.3490564**
3	-.0461775	**.0711808**	**-0.65**	**0.517**	**-.1856893**	**.0933342**
Wealth index combined						
2	.0007165	**.0955584**	**0.01**	**0.994**	**-.1865744**	**.1880074**
3	.2136284	**.1023026**	**2.09**	**0.037**	**.013119**	**.4141378**
4	.4291086	**.1102027**	**3.89**	**0.000**	**.2131153**	**.645102**
5	.8591041	**.1481702**	**5.80**	**0.000**	**.568696**	**1.149512**
Source of drinking water						
	-.0832925	**0646406**	**-1.29**	**0.198**	**-.2099857**	**.0434007**
Type of cooking fuel						
	-.5983471	**.08398**	**-7.12**	**0.000**	**-.7629449**	**-.4337493**
Pregnancy status						
	.0225564	**.0949293**	**0.24**	**0.812**	**-.1635016**	**.2086145**
Number of living children						
	-.0793552	**.0164873**	**-4.81**	**0.000**	**-.1116697**	**-.0470407**
Use of contraception						
	.0135541	**.0668866**	**0.20**	**0.839**	**-.1175412**	**.1446495**
Marital status						
2	.4416944	**.0823548**	**-5.36**	**0.000**	**-.6031068**	**-.2802821**
3	-.6756758	**.1134878**	**-5.95**	**0.000**	**-.8981077**	**-.4532439**

4	-.5816616	**.1753367**	**-3.32**	**0.001**	**-.9253153**	**-.2380079**
Availability of electricity						
	-2168265	**.0956523**	**-2.27**	**0.023**	**-.4043016**	**-.0293514**
Household has radio						
	.0284235	**.0567623**	**0.50**	**0.617**	**-.0828286**	**.1396757**
Household has Tv						
	.1243206	**.1048779**	**1.19**	**0.236**	**-.0812364**	**.3298776**
Household has refrigator						
	.2877088	**.0890919**	**3.23**	**0.001**	**.1130919**	**.4623258**
/cut1	-3.654924	**.1392851**			**-3.927918**	**-3.38193**
/cut2	.9844308	**.1333864**			**.7229983**	**1.245863**
/cut3	1.869218	**.134782**			**1.60505**	**2.133386**
Region Var(_cons)	.0180146	**.0110347**			**.0054229**	**.059844**
LR test vs. ologit model: ***chibar2(01)= 13.17*** ***Prob >= chibar2 = 0.0001***						